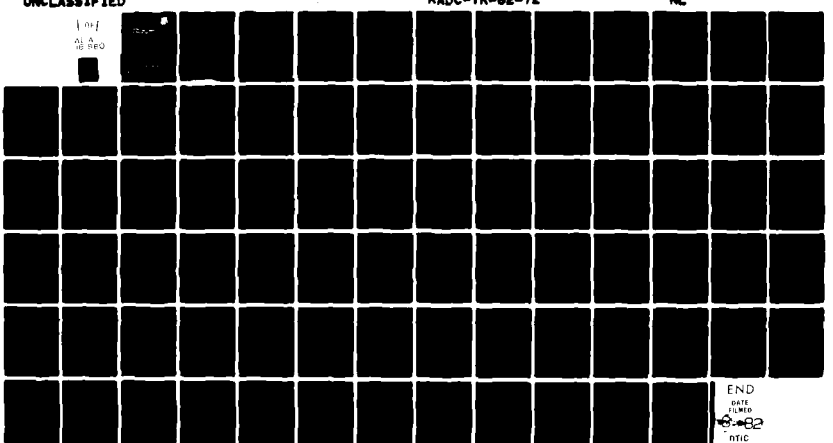
AD-A116 860 SOUTHERN ILLINOIS UNIV CARBONDALE F/G 12/1
PREPROCESSING TECHNIQUES IN TRANSIENT ANALYSIS.(U)
APR 82 D W KAMMLER F30602-78-C-0148
UNCLASSIFIED RADC-TR-82-72 NL



END
DATE FILMED
8-82
DTIC

(12)

RADC-TR-82-72
Final Technical Report
April 1982

AD A116860

PREPROCESSING TECHNIQUES IN TRANSIENT ANALYSIS

Southern Illinois University at Carbondale

David W. Kammler

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

This effort was funded totally by the Laboratory Directors' Fund

DTIC FILE COPY

ROME AIR DEVELOPMENT CENTER
Air Force Systems Command
Griffiss Air Force Base, NY 13441

DTIC
ELECTE
JUL 13 1982
E

82 07 13 003

[redacted]

MISSION
of
Rome Air Development Center

RADC plans and executes research, development, test and selected acquisition programs in support of Command, Control Communications and Intelligence (C^3I) activities. Technical and engineering support within areas of technical competence is provided to ESD Program Offices (POs) and other ESD elements. The principal technical mission areas are communications, electromagnetic guidance and control, surveillance of ground and aerospace objects, intelligence data collection and handling, information system technology, ionospheric propagation, solid state sciences, microwave physics and electronic reliability, maintainability and compatibility.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER RADC-TR-82-72	2. GOVT ACCESSION NO. AD-A116860	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) PREPROCESSING TECHNIQUES IN TRANSIENT ANALYSIS		5. TYPE OF REPORT & PERIOD COVERED Final Technical Report 6. PERFORMING ORG. REPORT NUMBER N/A
7. AUTHOR(s) David W. Kammler		8. CONTRACT OR GRANT NUMBER(s) F30602-78-C-0148
9. PERFORMING ORGANIZATION NAME AND ADDRESS Southern Illinois University at Carbondale Carbondale IL 62901		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS LD1204P1
11. CONTROLLING OFFICE NAME AND ADDRESS Rome Air Development Center (COT) Griffiss AFB NY 13441		12. REPORT DATE April 1982 13. NUMBER OF PAGES 73
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) Same		15. SECURITY CLASS. (of this report) UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE N/A

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

Same

18. SUPPLEMENTARY NOTES

RADC Project Engineer: Haywood E. Webb (COT)

This effort was funded totally by the Laboratory Directors' Fund.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Prony Algorithm
Dirichlet Series
Pole Location
Decomposition

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

In the Prony algorithm one determines pole locations from uniformly spaced wave form samples. In practice those samples are noisy which limits the effectiveness of the algorithm. This report first subjects the data samples to a pole invarient linear transformation to reduce the signal to noise ratio. The particular linear transformation chosen is based upon apriori knowledge of the regions where the poles are located.

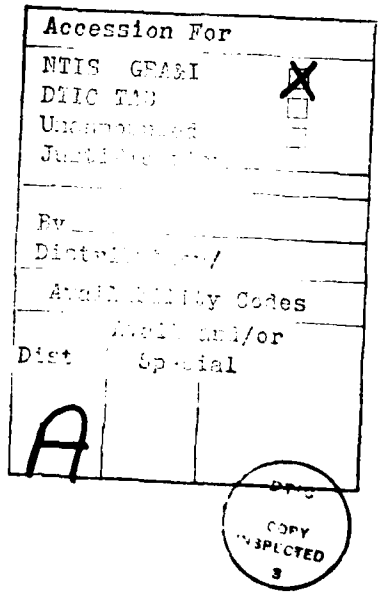
DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

ABSTRACT

Transient analysis deals with the problem of numerically determining the poles $s_1, s_2, \ldots$ associated with the Laplace transform of a given real transient $y(t) = \Sigma\, a_\nu \exp(s_\nu t)$ (where $\mathrm{Re}\, s_\nu < 0$ for each ν), from a knowledge of its samples $y_k = y(k\,\Delta t)$, $k = 0,1,\ldots$. In practice, the samples $y_0, y_1, y_2, \ldots$ are usually contaminated with noise, and this serves to limit the effectiveness of the computational schemes of Prony, Bellman, Jain, Van Blaricum, etc. which have been developed for extracting the first few s_ν's. The performance of these algorithms can be greatly enhanced if the data $y_0, y_1, y_2, \ldots$ is first subjected to a suitable sequence-to-sequence transformation. Any such linear pole preserving transformation must have the simple form $w_k = \gamma_0 y_k + \gamma_1 y_{k+1} + \gamma_2 y_{k+2} + \ldots$, $k = 0,1,\ldots$ where $\gamma(z) = \gamma_0 + \gamma_1 z + \gamma_2 z^2 + \ldots$ is analytic and zero free in the unit disc $|z| < 1$. In most cases of interest, $\gamma(z)$ may be chosen so as to greatly suppress the effects of noise.



PREPROCESSING TECHNIQUES IN TRANSIENT ANALYSIS

TABLE OF CONTENTS

ABSTRACT 1

1. INTRODUCTION TO TRANSIENT ANALYSIS 2

2. PREPROCESSING WITH THE TAIL SUM OPERATOR 10

3. PREPROCESSING OPERATORS 15

4. NOISE SUPPRESSION 27

5. CONSTRUCTING NOISE SUPPRESSING PREPROCESSING OPERATORS 36

6. ALTERNATIVE METHODS FOR COMPUTING THE NULL VECTOR . 41

7. THE PROBLEM OF EXTRANEOUS ROOTS 52

8. CONCLUSIONS 63

9. APPENDIX - THE DAMPED VIBRATING STRING 65

FIGURES 69

REFERENCES 73

PREPROCESSING TECHNIQUES IN TRANSIENT ANALYSIS

David W. Kammler

ABSTRACT

We consider the problem of finding the poles $z_1, z_2, \ldots, z_n$ associated with the z- transform of the sequence $\mathbf{y} = (y_0, y_1, y_2 \ldots)$ of samples

$$y_k = Y(kT), \quad k = 0,1,2,\ldots$$

of the transient

$$Y(t) = \sum_{\nu=1}^{n} a_\nu \exp(s_\nu t), \quad t > 0$$

emitted by a given n-th order linear system. In principle, exactly the same poles can be extracted from the sequence $\mathbf{u} = \Gamma \mathbf{y}$ when Γ is a sequence-to-sequence transformation of the form $\Gamma = \gamma(E)$ where E is the shift operator and $\gamma(z)$ is analytic and zero free on the unit disc $|z| < 1$. Such a preprocessing operator Γ can be chosen so as to suppress additive noise or to selectively enhance one or more of the poles without annihilating the others. Using such preprocessing operators we obtain a common conceptual framework for all of the previously used schemes for transient analysis (including those of Prony, Van Blaricum & Mittra, and Jain) and we provide a theoretical basis for several promising new algorithms.

1. INTRODUCTION TO TRANSIENT ANALYSIS

Let

$$(1)\quad Y(t) = \sum_{\nu=1}^{n} a_\nu e^{s_\nu t},\ t \geq 0$$

denote an n-th order approximation to the transient emitted by a physical system in response to some initial excitation cf.[1]. (In the appendix we present a mathematical model for a damped vibrating string which serves to illustrate the kind of phenomena we wish to analyze.) We consider the problem of numerically determining n and the complex frequencies $s_1,\ldots,s_n$ from a knowledge of the sequence $\mathbf{y}=(y_0,y_1,y_2,\ldots)$ of uniformly spaced samples

$$(2)\quad y_k = Y(kT),\ k = 0,1,2,\ldots$$

where $T > 0$ is the sampling interval. We shall assume that Y is real valued so that the a_ν's and s_ν's occur in complex conjugate pairs, and we assume that Y is nondegenerate in the sense that the s_ν's are distinct and the a_ν's are nonzero. We further assume that the s_ν's all lie in the left half plane so that $Y(t)\to 0$ as $t \to \infty$. By substituting $t=kT$ in (1) we see that the samples have the representation

$$(3)\quad y_k = \sum_{\nu=1}^{n} a_\nu z_\nu^k,\ k=0,1,2,\ldots$$

where

$$(4)\quad z_\nu = e^{s_\nu T},\ \nu=1,\ldots,n$$

with

(5) $|z_\nu| < 1, \ \nu=1,\ldots,n.$

Since $s_1,\ldots,s_n$ occur in complex conjugate pairs the same is true of $z_1,\ldots,z_n$ so that the coefficients $c_0,c_1,\ldots,c_n$ of the corresponding characteristic polynomial

(6) $c(z) = c_0 + c_1 z + \ldots + c_n z^n = c_n(z-z_1)\ldots(z-z_n), c_n \in \mathbb{R}, c_n \neq 0,$

are all real valued. We shall assume that the sampling interval T is sufficiently small so that $z_1,\ldots,z_n$ are distinct and so that the inversion of (4)

(7) $s_\nu = T^{-1} \ell n z_\nu, \ \nu=1,\ldots,n$

can be effected unambiguously without undue regard to aliasing.

We shall find it convenient to introduce the shift operator E which is defined so that

(8) $E(y_0,y_1,y_2,\ldots) = (y_1,y_2,y_3,\ldots)$

or equivalently

$$Ey_k = y_{k+1}, \ k=0,1,2,\ldots,$$

and we shall let I denote the corresponding identity operator. We observe that

$$E(1,z,z^2,\ldots) = z\cdot(1,z,z^2,\ldots)$$

so that the application of the operator E-zI annihilates the power sequence $(1,z,z^2,\ldots)$. This being the case, since (3) expresses $\mathbf{y}$ as a linear combination of n such power sequences corresponding to $z_1,\ldots,z_n$ we must have

$$(9)\quad (E-z_1)\ldots(E-z_n)\mathbf{y} = \mathbf{0}.$$

By using (6) we may rewrite (9) in the form

$$c(E)\mathbf{y}=(c_0I+c_1E+\ldots+c_nE^n)\mathbf{y}=c_0\mathbf{y}+c_1(E\mathbf{y})+\ldots+c_n(E^n\mathbf{y})=\mathbf{0}$$

or equivalently

$$(10)\quad \mathcal{Y}\,\mathbf{c} = \mathbf{0}$$

where the data matrix

$$(11)\quad \mathcal{Y} = \begin{bmatrix} y_0 & y_1 & \cdots & y_n \\ y_1 & y_2 & \cdots & y_{n+1} \\ \vdots & \vdots & & \vdots \end{bmatrix}$$

has the columns $\mathbf{y}$, $E\mathbf{y},\ldots,$ $E^n\mathbf{y}$ and where $\mathbf{c} = (c_0,c_1,\ldots,c_n)^T$. Since we have assumed that the a_ν's are nonzero and the z_ν's are distinct, the first n columns of $\mathcal{Y}$ are linearly independent, $\mathcal{Y}$ has rank n, and (apart from a scale factor) the null vector $\mathbf{c}$ is uniquely determined.

These observations suggest the following approach to the transient analysis problem. Given the samples $\mathbf{y} = (y_0,y_1,y_2,\ldots)$ we attempt to make some slight overestimate of the system order n and form the data matrix (11). We numerically investigate the null space of $\mathcal{Y}$ and, if

necessary, reduce n so as to obtain an essentially unique null vector **c**. After effecting the factorization (6) and using (7) we then obtain the desired s_ν's. Almost all of the presently used noniterative schemes for computing the s_ν's from the y_k's fall within this conceptual framework. For example, when we use the well known method of Prony, we ignore all but the first n rows of $\mathcal{Y}$ and after arbitrarily setting $c_n=1$ we solve the resulting system of linear equations

$$(12)\quad \begin{bmatrix} y_0 & y_1 & \cdots & y_n \\ y_1 & y_2 & \cdots & y_{n+1} \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ y_{n-1} & y_n & \cdots & y_{2n-1} \end{bmatrix} \begin{bmatrix} c_0 \\ \cdot \\ \cdot \\ \cdot \\ c_{n-1} \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix}$$

to obtain the remaining components of the approximate null vector **c**.

Perhaps the most natural approach to finding a null vector for $\mathcal{Y}$ (especially when $\mathcal{Y}$ has been contaminated with noise) is to determine **c** so as to minimize the Euclidean length of $\mathcal{Y}\mathbf{c}$ subject to some normalization of the Euclidean length of **c**, or equivalently, to minimize the Rayleigh quotient

$$(13)\quad \frac{\mathbf{c}^T\mathcal{Y}^T\mathcal{Y}\mathbf{c}}{\mathbf{c}^T\mathbf{c}} = \frac{\|\mathcal{Y}\mathbf{c}\|_2^2}{\|\mathbf{c}\|_2^2}$$

(In so doing we avoid changing an intrinsically homogeneous

problem into an inhomogeneous one as in (12) by the imposition of a constraint $c_n=1$ on one of the components of **c**.) Any minimizing **c** for (13) must be an eigenvector belonging to the smallest eigenvalue of the symmetric nonnegative semidefinite (n+1) x (n+1) matrix $\mathcal{Y}^T\mathcal{Y}$, cf.[8,p.266]. In the ideal noise free setting where the system order n is known, both $\mathcal{Y}$ and $\mathcal{Y}^T\mathcal{Y}$ will have rank n, the minimum eigenvalue of $\mathcal{Y}^T\mathcal{Y}$ will be zero, and the corresponding eigenvector **c** (unique to within a scale factor) will be the desired null vector of $\mathcal{Y}$. In practice, we do not know the system order and we can only approximate the elements

$$(14)\quad (\mathcal{Y}^T\mathcal{Y})_{ij} = \sum_{k=0}^{\infty} y_{i+k}\, y_{j+k},\ i,j, = 0,1,\ldots,\ n$$

due to noise contamination of the data and to the use of finite precision arithmetic on our computer so our computed $\mathcal{Y}^T\mathcal{Y}$ will not have either a zero eigenvalue or a null vector. Nevertheless, by using well known numerical methods [7] we can effect the eigenvalue decomposition

$$(15)\quad \mathcal{Y}^T\mathcal{Y} = \lambda_0\mathbf{v}_0\mathbf{v}_0^T + \lambda_1\mathbf{v}_1\mathbf{v}_1^T + \ldots + \lambda_n\mathbf{v}_n\mathbf{v}_n^T$$

where $\lambda_0 \geq \lambda_1 \geq \ldots > \lambda_n \geq 0$ are the eigenvalues and $\mathbf{v}_0,\mathbf{v}_1,\ldots,\mathbf{v}_n$ a corresponding orthonormal set of eigenvectors for $\mathcal{Y}^T\mathcal{Y}$. If we purposely use a value of n which is a bit too large, we can analyze the distribution of the λ_k's and thereby ascertain the correct system order, e.g., when n is

correctly chosen we might expect to have $\lambda_{n-1} >> \lambda_n$ and $\lambda_n \approx 0$. The corresponding $\mathbf{v}_n$ is then a good choice for $\mathbf{c}$. This approach has been successfully developed by VanBlaricum and his coworkers, [1,9].

As an alternative to the eigenvalue decomposition (15), we can use a singular value decomposition of $\mathcal{Y}$ to obtain the $\mathbf{c}$ which minimizes (13). The data matrix $\mathcal{Y}$ can be represented in the form

$$(16)\quad \mathcal{Y} = \sigma_0 \mathbf{u}_0 \mathbf{v}_0^T + \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \ldots + \sigma_n \mathbf{u}_n \mathbf{v}_n^T$$

where $\sigma_0 \geq \sigma_1 \geq \ldots \geq \sigma_n \geq 0$ are the singular values of $\mathcal{Y}$, where $\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_n$ are orthonormal vectors from $\mathbb{R}^{n+1}$, and where $\mathbf{u}_0, \mathbf{u}_1, \ldots, \mathbf{u}_n$ are orthonormal sequences. (Indeed, $\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_n$ are again an orthonormal system of eigenvectors of $\mathcal{Y}^T\mathcal{Y}$ as in (15) with $\sigma_k = \lambda_k^{1/2}$, k=0,1,..., n and $\mathbf{u}_k = \sigma_k^{-1} \mathcal{Y} \mathbf{v}_k$ for each k for which $\sigma_k \neq 0$.) The σ_k's and $\mathbf{v}_k$'s can be computed directly from the elements of the data matrix $\mathcal{Y}$ without first forming $\mathcal{Y}^T\mathcal{Y}$, cf.[2,3]. The distribution of the σ_k's can be used in the same way as that of the λ_k's to help determine the system order, and the right singular vector $\mathbf{v}_n$ serves as an approximate null vector $\mathbf{c}$ for $\mathcal{Y}$.

In principle, the singular value decomposition (16) should result in a somewhat better conditioned estimate of the approximate null vector $\mathbf{c}$ than the eigenvalue decomposition (15). Assuming $\sigma_{n-1} > \sigma_n$, it can be shown that when $\mathcal{Y}$

is replaced by $\mathcal{Y} + \mathcal{E}$ in (16) the resulting perturbation $\Delta\mathbf{c}$ which results in the normalized approximate null vector $\mathbf{c} = \mathbf{v}_n$ satisfies the bound

$$(17) \quad \frac{\| \Delta\mathbf{c} \|_2}{\| \mathbf{c} \|_2} \leq \frac{\sigma_0}{\sigma_{n-1} - \sigma_n} \cdot \| \mathcal{E} \|_2 + o(\| \mathcal{E} \|_2)$$

whereas when $\mathcal{Y}^T\mathcal{Y}$ is replaced by $\mathcal{Y}^T\mathcal{Y} + \mathcal{W}$ in (15) we have the corresponding bound

$$(18) \quad \frac{\| \Delta\mathbf{c} \|_2}{\| \mathbf{c} \|_2} \leq \frac{\lambda_0}{\lambda_{n-1} - \lambda_n} \cdot \| \mathcal{W} \|_2 + o(\| \mathcal{W} \|_2),$$

with both bounds being sharp. The condition number

$$(19) \quad \kappa_E = \frac{\lambda_0}{\lambda_{n-1} - \lambda_n} = \frac{\sigma_0^2}{\sigma_{n-1}^2 - \sigma_n^2}$$

associated with the null vector computation based on (15) is usually many orders of magnitude larger than the corresponding condition number

$$(20) \quad \kappa_S = \frac{\sigma_0}{\sigma_{n-1} - \sigma_n}$$

associated with the null vector computation based on (16). On the other hand, when using the eigenvalue decomposition we must store only one copy of (a suitable truncation of) the data sequence $\mathbf{y} = (y_0, y_1, y_2, \ldots)$ to use in computing the elements (14) of $\mathcal{Y}^T\mathcal{Y}$, and this symmetric (n+1)x(n+1) matrix is then used as the input to a routine which performs

the eigenvalue computations. In contrast, we must assemble $\mathcal{Y}$ (which requires n+1 times as much storage as $\mathbf{y}$) to use as the input for a singular value decomposition analysis. When the system order n is large and when many rows of $\mathcal{Y}$ are known and available for use, this additional storage requirement for the singular value decomposition may very well prohibit its use.

The above formulation of the transient analysis problem has been strongly influenced by a recent paper of Henderson [4] and by Volume I of the technical report [1] of Auton and Van Blaricum. An unusually complete annotated bibliography of related papers and technical reports is given in Volume III of [1].

2. PREPROCESSING WITH THE TAIL SUM OPERATOR

Suppose that we are given a transient sequence $\mathbf{y} = (y_0, y_1, y_2, \ldots)$ having components (3) with $a_1, \ldots, a_n$ being nonzero and with $z_1, \ldots, z_n$ being distinct points within the unit circle, and that we wish to determine the z_ν's numerically. Although it is possible to process tne y_k's directly as described in the previous section, it is often advantageous to use the y_k's to generate an auxiliary sequence $\mathbf{u} = (u_0, u_1, u_2 \ldots)$ from which we subsequently extract the z_ν's . This procedure is known as preprocessing. Before giving a more precise definition, we shall consider a specific example.

We define the tail sum operator S to be the sequence-to-sequence mapping for which

$$(21) \quad Sy_k = \sum_{\ell=0}^{\infty} y_{k+\ell} \ , \ k=0,1,2,\ldots \ ,$$

i.e., to generate the elements of $\mathbf{u} = S\mathbf{y}$ we compute

$$(22) \quad u_k = y_k + y_{k+1} + y_{k+2} + \ldots \ , \ k=0,1,2,\ldots$$

directly from the elements of $\mathbf{y}$. Using (21) we see that

$$Sz^k = \sum_{\ell=0}^{\infty} z^{k+\ell} = z^k/(1-z)$$

so that

$$S(1, z, z^2, \ldots) = (1-z)^{-1} \cdot (1, z, z^2, \ldots) \ ,$$

i.e., the power sequence $(1,z,z^2,\ldots)$ is an eigenvector of S corresponding to the eigenvalue $(1-z)^{-1}$ for every choice of $|z| < 1$. This being the case, if **y** has the representation (3) (so that **y** is a linear combination of n such power sequences), then **u**=S**y** has the components

$$(23) \quad u_k = \sum_{\nu=1}^{n} a_\nu^* z_\nu^k , \quad k=0,1,2,\ldots$$

where

$$(24) \quad a_\nu^* = a_\nu/(1-z_\nu), \quad \nu=1,\ldots, n.$$

We thereby see that the sequence **u**=S**y** has exactly the same poles $z_1,\ldots,z_n$ as **y**, so it is possible to extract these poles from **u** as described in the previous section. Using (24) we see that the ratio a_ν^*/a_ν is large when z_ν is near 1, and thus the pole z_ν is more strongly represented in **u** than in **y** when this is the case. Moreover, we might expect the summation (22) to suppress some of the effects of any noise which may have contaminated the y_k's. This being the case, we might reasonably hope to extract slightly more accurate poles from **u** than we could obtain directly from **y**.

Of course, if a single application of S tends to suppress the noise and enhance the poles which lie near z=1, the repeated application of this operator might very well be expected to do an even better job. The application of S to the sequence S**y** gives the sequence S^2**y**, etc., with each of

the sequences $\mathbf{y}$, $S\mathbf{y}$, $S^2\mathbf{y}$,... having exactly the same poles $z_1,\ldots,z_n$. This suggests three different schemes for computing these poles. First of all, we might select some fixed p=0,1,2,... and, following our earlier approach, attempt to find a null vector $\mathbf{c}$ for the data matrix

$$(25)\qquad \begin{bmatrix} S^p y_0 & S^p y_1 & \cdots & S^p y_n \\ S^p y_1 & S^p y_2 & \cdots & S^p y_{n+1} \\ S^p y_2 & S^p y_3 & \cdots & S^p y_{n+2} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \end{bmatrix}$$

(which reduces to (11) when we take p=0 and $S^0 = I$.) Upon effecting the corresponding factorization (6) we obtain the z_ν's . The second approach is suggested by the observation that in the absence of noise the exact null vector $\mathbf{c}$ is orthogonal to the first row of the matrix (25) for each p=0,1,2,... . We might therefore seek a null vector $\mathbf{c}$ for the matrix

$$(26)\qquad \begin{bmatrix} y_0 & y_1 & \cdots & y_n \\ Sy_0 & Sy_1 & \cdots & Sy_n \\ S^2 y_0 & S^2 y_1 & \cdots & S^2 y_n \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \end{bmatrix}$$

and again obtain the z_ν's from the resulting factorization (6). Finally, by repeatedly applying S to (3) we see that

$$(27)\qquad S^p y_k = \sum_{\nu=1}^{n} a_\nu z_\nu^k \left[(1-z_\nu)^{-1}\right]^p, \quad p=0,1,2,\ldots, \quad k=0,1,2,\ldots$$

and thereby observe that for each fixed k=0,1,2,... the sequence y_k , Sy_k, S^2y_k,... has the associated poles $w_\nu = (1-z_\nu)^{-1}$, $\nu=1,\ldots,n$ (which lie in the half plane Re w > 1/2 when $|z_\nu| < 1$ for each ν.) This being the case, if we find a null vector $\mathbf{d} = (d_0,d_1\ldots,d_n)^T$ for the matrix

$$(28)\quad \begin{bmatrix} y_0 & Sy_0 & \ldots & S^n y_0 \\ y_1 & Sy_1 & \ldots & S^n y_1 \\ y_2 & Sy_2 & \ldots & S^n y_2 \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \end{bmatrix}$$

and effect the corresponding factorization

$$d_0 + d_1 w + \ldots + d_n w^n = d_n(w-w_1) \ldots (w-w_n)$$

we will have $w_\nu = (1-z_\nu)^{-1}$ or equivalently $z_\nu = (w_\nu-1)/w_\nu$, $\nu=1,\ldots,n$ (after a suitable permutation of the indices.) This third approach is equivalent to the pencil-of-functions method of Jain and his coworkers, [5,6].

If **y** has the exact representation (3) and all computations are performed without error, these three approaches all yield the same z_ν's. In practice, however, the system order may be infinite (with n of the poles being dominant and the influence of the others being small but nonzero), the sampled y_k's may be contaminated with noise, and finite precision arithmetic is used to carry out the computations. The accuracy of the computed z_ν's thus depends on which of

the three matrices (25), (26), (28) we use and on which method we use to determine the approximate null vector.

3. PREPROCESSING OPERATORS

We shall now generalize the results of the previous section so as to include preprocessing schemes other than those based on the tail sum operator S. Let

$$\mathbf{u} = \Gamma \mathbf{y} \tag{29}$$

be a linear sequence-to-sequence transformation. We may think of $\mathbf{u},\mathbf{y}$ as being (column) vectors with $\Gamma = [\gamma_{k\ell}]$ being a matrix so that

$$u_k = \Gamma y_k = \gamma_{k0}y_0 + \gamma_{k1}y_1 + \gamma_{k2}y_2 + \ldots \quad . \tag{30}$$

We would like to impose restrictions on the $\gamma_{k\ell}$'s which will insure that Γ is defined on the whole set of rapidly decaying transient sequences

$$\mathcal{T} = \{ (y_0,y_1,y_2,\ldots)^T : \limsup |y_k|^{1/k} < 1 \} \quad . \tag{31}$$

(When y_k has the form (3) with $a_\nu \neq 0$ and $|z_\nu| < 1$ for each $\nu=1,\ldots,n$ we find

$$\limsup |y_k|^{1/k} = \max \{ |z_1|,\ldots,|z_n| \} < 1$$

so that $\mathcal{T}$ includes all of the transient sequences which might arise from any stable finite order system we might wish to study.) Moreover, to be useful as a preprocessing scheme we must insist that Γ be pole preserving in the sense that z_ν is a system pole associated with $\mathbf{u} = \Gamma \mathbf{y}$ if and only

if z_ν is also a system pole associated with **y**, and this serves to further restrict the $\gamma_{k\ell}$'s. The resulting sequence-to-sequence mappings are characterized by the following

THEOREM 1: A necessary and sufficient condition for Γ to be a linear pole preserving sequence-to-sequence mapping of $\mathcal{Y}$ into $\mathcal{Y}$ is that Γ have the form

$$(32)\quad \Gamma = \gamma(E) = \gamma_0 + \gamma_1 E + \gamma_2 E^2 + \ldots$$

i.e.,

$$(33)\quad \Gamma y_k = \gamma_0 y_k + \gamma_1 y_{k+1} + \gamma_2 y_{k+2} + \ldots\ , \quad k=0,1,2,\ldots$$

where

$$(34)\quad \gamma(z) = \gamma_0 + \gamma_1 z + \gamma_2 z^2 + \ldots$$

is a zero free analytic function on the unit disc $|z| < 1$.

Before proving this result, we point out that (32)-(33) imply that Γ has the banded upper triangular matrix representation

$$(35)\quad \Gamma = \begin{bmatrix} \gamma_0 & \gamma_1 & \gamma_2 & \gamma_3 & \cdots \\ & \gamma_0 & \gamma_1 & \gamma_2 & \cdots \\ & & \gamma_0 & \gamma_1 & \cdots \\ & & & \gamma_0 & \cdots \\ & & & & \ddots \end{bmatrix}$$

and since the radius of convergence, R, of (34) is given by

the Cauchy-Hadamard formula

$$(36)\quad R^{-1} = \limsup |\gamma_k|^{1/k} ,$$

the analyticity of $\gamma(z)$ on $|z| < 1$ is equivalent to the requirement

$$(37)\quad \limsup |\gamma_k|^{1/k} \leq 1 .$$

The hypothesis that $\gamma(z)$ be zero free is not easily translated into a simple condition on the γ_k's .

To illustrate the theorem, we first note that the tail sum operator S of (22) has the representation

$$(38)\quad S = \gamma(E) = I + E + E^2 + \ldots$$

where

$$(39)\quad \gamma(z) = 1+z+z^2+ \ldots = 1/(1-z)$$

is clearly analytic and zero free on $|z| < 1$. Likewise, the local smoothing scheme

$$(40)\quad u_k = y_k + y_{k+1} + \ldots + y_{k+N-1}$$

which results from the operator

$$(41)\quad S_N = \gamma_N(E) = I + E + E^2 + \ldots + E^{N-1}$$

with

$$(42)\quad \gamma_N(z) = 1 + z + z^2 = \ldots + z^{N-1} = (1-z^N)/(1-z)$$

meets the hypotheses of the theorem. On the other hand, the weighted smoothing scheme

$$u_k = y_k + 2y_{k+1} + 3y_{k+2}$$

corresponds to the operator

(43) $\Gamma = \gamma(E) = I + 2E + 3E^2$

which maps $\mathcal{T}$ into $\mathcal{T}$, but since

$$\gamma(z) = 1 + 2z + 3z^2$$

has the roots $(-1 \pm \sqrt{-2}\,)/3$ which lie in the unit disc, this map is not always pole-preserving.

We shall now state and prove three lemmas which collectively serve to establish the above theorem. The first focuses on the banded upper triangular structure (35) of Γ, the second on the growth condition (37), and the third on the requirement that $\gamma(z)$ be zero free in the unit disc.

LEMMA 1. Let Γ be a linear operator which maps the space of transients $\mathcal{T}$ into itself. The following are equivalent:

(i) $c(E)\Gamma \mathbf{y} = \mathbf{0}$ whenever $\mathbf{y} \in \mathcal{T}$ and $c(z)$ is a polynomial such that $c(E)\mathbf{y} = \mathbf{0}$, (i.e., if y_k has the representation (3) for some choice of the a_ν's and z_ν's, then $u_k = \Gamma y_k$ must have the repre-

sentation (23) for some choice of the a_ν^*'s with the z_ν's being the same.),

(ii) Γ has the representation of (32)-(35),

(iii) Γ commutes with the shift operator E.

Proof. Assuming (i) we see that since

$$(E - zI)\,(1, z, z^2, \ldots) = (0, 0, 0, \ldots)$$

we must also have

$$(E - zI)\,\Gamma\,(1, z, z^2, \ldots) = (0, 0, 0, \ldots)$$

for each choice of z with $|z| < 1$. Using the general representation (30) for the linear operator Γ we find

$$(E - zI)\,\Gamma\,(1, z, z^2, \ldots)$$

$$= (E - zI)\left(\sum_{\ell=0}^{\infty} \gamma_{0\ell} z^\ell,\ \sum_{\ell=0}^{\infty} \gamma_{1\ell} z^\ell,\ \sum_{\ell=0}^{\infty} \gamma_{2\ell} z^\ell, \ldots\right)$$

$$= \left(\gamma_{10} + \sum_{\ell=1}^{\infty} (\gamma_{1\ell} - \gamma_{0,\ell-1}) z^\ell,\ \gamma_{20} + \sum_{\ell=1}^{\infty} (\gamma_{2\ell} - \gamma_{1,\ell-1}) z^\ell, \ldots\right)$$

so that the power series

$$\gamma_{k0} + \sum_{\ell=1}^{\infty} (\gamma_{k\ell} - \gamma_{k-1,\ell-1}) z^\ell, \quad k = 1, 2, \ldots$$

must vanish for every choice of z with $|z| < 1$. It follows that

$$\gamma_{k0} = 0, \quad k = 1,2,\ldots$$

$$\gamma_{k\ell} = \gamma_{k-1,\ell-1} \ , \quad k,\ell = 1,2,\ldots \ ,$$

i.e., that Γ has the representation (35) or equivalently (32) so that (i) implies (ii).

When $\Gamma = \gamma(E)$ has the representation of (32)-(33) we clearly have

$$\Gamma E y_k = \Gamma y_{k+1} = \gamma_0 y_{k+1} + \gamma_1 y_{k+2} + \gamma_2 y_{k+3} + \ldots = E\Gamma y_k, k=0,1,\ldots$$

for each $\mathbf{y} \in \mathcal{T}$ so that (ii) implies (iii). Moreover, if Γ commutes with E, then Γ also commutes with the polynomial $c(E)$ in E so that

$$c(E)\ \Gamma\ \mathbf{y} = \Gamma\, c(E)\mathbf{y} = \Gamma\, \mathbf{0} = \mathbf{0}$$

whenever $c(E)\ \mathbf{y} = \mathbf{0}$, i.e., (iii) implies (i). □

LEMMA 2. Let Γ be an operator having the representation of (32) - (33). In order that Γ map $\mathcal{T}$ into $\mathcal{T}$ it is both necessary and sufficient that (37) hold (or equivalently, that (34) be analytic on the disc $|z| < 1$.)

Proof. Assume first that Γ maps $\mathcal{T}$ into $\mathcal{T}$. The series (34) which represents the first component of $\Gamma(1,z,z^2,\ldots)$ must then converge whenever $(1,z,z^2,\ldots) \in \mathcal{T}$, i.e., whenever $|z| < 1$. Thus (34) is analytic on the unit disc so that the Cauchy-Hadamard formula (36) must produce a radius of convergence $R \geq 1$, and (37) holds.

Conversely, assume that (37) holds and that $\mathbf{y} \in \mathcal{T}$ is given so that

(44) $\limsup |y_k|^{1/k} < 1$.

Since (44) holds, there exist constants $A > 0$, $0 < \alpha < 1$ such that

(45) $|y_k| \leq A\ \alpha^k$, $k=0,1,2,\ldots$,

and since (37) also holds there exist constants $B > 0$, $0 < \beta < \alpha^{-1}$ such that

(46) $|\gamma_k| \leq B\ \beta^k$, $k=0,1,2,\ldots$.

Using (45) - (46) and the fact that $0 < \alpha\beta < 1$ we see that the series (33) which is used for the k-th component of $\Gamma\mathbf{y}$ is majorized by

$$|\Gamma y_k| = |\gamma_0 y_k + \gamma_1 y_{k+1} + \gamma_2 y_{k+2} + \ldots|$$

$$\leq B\beta^0 \cdot A\alpha^k + B\beta^1 \cdot A\alpha^{k+1} + B\beta^2 \cdot A\alpha^{k+2} + \ldots$$

$$= AB\alpha^k / (1-\alpha\beta),$$

and is thus convergent, $k=0,1,2,\ldots$. Thus $\Gamma\mathbf{y}$ is well defined, and

$$\limsup |\Gamma y_k|^{1/k} \leq \limsup |AB\alpha^k/(1-\alpha\beta)|^{1/k} = \alpha < 1$$

so that $\Gamma\mathbf{y} \in \mathcal{T}$. □

NOTE. A slight extension of the above argument shows that the operator of (32) - (33) will map

$$\mathcal{T}_R = \left\{ (y_0, y_1, y_2, \dots)^T : \limsup |y_k|^{1/k} < R \right\}$$

into itself provided that

$$\limsup |\gamma_k|^{1/k} \leq 1/R,$$

i.e., provided $\gamma(z)$ is analytic on the disc $|z| < R$, $0 < R < \infty$. The case R=1 is covered by the lemma.

When $\gamma(z)$ is analytic on the unit disc and $|z| < 1$ we have

$$\gamma(E)\ (1, z, z^2, \dots) = \gamma(z)\ (1, z, z^2, \dots),$$

and we thereby see that $\gamma(E)$ will annihilate the power sequence associated with the pole $z=z_0$ if z_0 is a root of $\gamma(z)$. More generally, suppose that $\mathbf{y}$ has the representation

$$(47) \quad y_k = \sum_{\nu=1}^{\infty} a_\nu z_\nu^k, \quad k=0,1,2,\dots$$

where

$$(48) \quad \sum_{\nu=1}^{\infty} |a_\nu| < \infty$$

and where

$$(49) \quad |z_\nu| \leq R < 1 \text{ for each } \nu=1,2,\dots .$$

We then find

$$\limsup |y_k|^{1/k} \le \limsup \left\{ \sum_{\nu=1}^{\infty} |a_\nu| \, |z_\nu|^k \right\}^{1/k}$$

$$\le \limsup \left\{ R^k \sum_{\nu=1}^{\infty} |a_\nu| \right\}^{1/k}$$

$$= R < 1$$

so that $\mathbf{y} \in \mathcal{T}$. Moreover, when $\gamma(z)$ is analytic on the unit disc we have

$$\gamma(E)y_k = \sum_{\ell=0}^{\infty} \gamma_\ell \sum_{\nu=1}^{\infty} a_\nu z_\nu^{k+\ell} = \sum_{\nu=1}^{\infty} \gamma(z_\nu) a_\nu z_\nu^k \,, k=0,1,2,\ldots,$$

and we thereby conclude that each pole z_ν associated with $\mathbf{y}$ will also be a pole of $\mathbf{u} = \gamma(E)\mathbf{y}$ provided that $\gamma(z)$ has no zeros in the unit disc.

The above arguments show that when $\gamma(z)$ is analytic and zero free on the unit disc, then $\gamma(E)$ is a pole preserving mapping of the set of sequences $\mathbf{y}$ of the form (47)-(49) into itself. Although such transients are the ones most likely to be met in practice, for the sake of completeness we shall extend the argument so as to include the somewhat larger class $\mathcal{T}$. In so doing, we shall find it convenient to use the notation

$$y(z) = y_0 + y_1 z + y_2 z^2 + \ldots$$

for the generating function associated with a given sequence

$\mathbf{y} = (y_0, y_1, y_2, \ldots)$, with

$$y(z^{-1}) = y_0 + y_1 z^{-1} + y_2 z^{-2} + \ldots$$

being the corresponding z - transform. Of course, when $\mathbf{y}$ has the representation (47) - (49) we find

$$y(z^{-1}) = \sum_{k=0}^{\infty} z^{-k} \sum_{\nu=1}^{\infty} a_\nu z_\nu^k = z \cdot \sum_{\nu=1}^{\infty} a_\nu/(z-z_\nu)$$

so that $z_1, z_2, \ldots$ are the poles of $y(z^{-1})$. More generally, we shall say that the map Γ is pole preserving on $\mathcal{T}$ provided that for every choice of $\mathbf{y} \in \mathcal{T}$ the z - transform $u(z^{-1})$ of $\mathbf{u} = \Gamma \mathbf{y}$ has exactly the same poles as the z-transform $y(z^{-1})$ of $\mathbf{y}$. The relationship between these poles is made precise in the following

LEMMA 3. Let $\mathbf{y}$, $\boldsymbol{\gamma}$ be sequences with

(50) $\limsup |y_k|^{1/k} = R < 1$, $\limsup |\gamma_k|^{1/k} \leq 1$

and let $\mathbf{u} = \gamma(E)\,\mathbf{y}$, $\mathbf{v} = y(E)\boldsymbol{\gamma}$, i.e.,

$$u_k = \gamma_0 y_k + \gamma_1 y_{k+1} + \gamma_2 y_{k+2} + \ldots, \quad k=0,1,2,\ldots$$

$$v_k = y_0 \gamma_k + y_1 \gamma_{k+1} + y_2 \gamma_{k+2} + \ldots, \quad k=0,1,2,\ldots \quad .$$

Then the generating functions $v(z)$, $\gamma(z)$ are analytic for $|z| < 1$, the z-transforms $y(z^{-1})$, $u(z^{-1})$ are analytic for $|z| > R$, and the identity

$$u(z^{-1}) = \gamma(z) \cdot y(z^{-1}) - v(z) + v(0)$$

holds in the annulus $R < |z| < 1$ and thus in the common domain of analyticity of these four functions. In particular, when γ is zero free in the unit disc, the poles of $u(z^{-1})$ (which must lie in the disc $|z| \leq R$) must coincide with the poles of $y(z^{-1})$.

Proof. By using the Cauchy-Hadamard formula in conjunction with (50) and the note following Lemma 2 we infer that $y(z)$, $u(z)$ are analytic for $|z| < 1/R$ and that $\gamma(z)$, $v(z)$ are analytic for $|z| < 1$. For z in the annulus $R < |z| < 1$ we then have

$$\gamma(z)\cdot y(z^{-1})+v(0)= \sum_{k=0}^{\infty}\sum_{\ell=0}^{\infty} \gamma_k z^k \cdot y_\ell z^{-\ell} + \sum_{k=0}^{\infty} \gamma_k y_k$$

$$= \sum_{k\geq\ell} \gamma_k y_\ell z^{k-\ell} + \sum_{\ell\geq k} \gamma_k y_\ell z^{-(\ell-k)}$$

$$= \sum_{\nu=0}^{\infty}\Big(\sum_{\ell=0}^{\infty} \gamma_{\ell+\nu} y_\ell\Big) z^{\nu} + \sum_{\nu=0}^{\infty}\Big(\sum_{k=0}^{\infty} \gamma_k y_{k+\nu}\Big) z^{-\nu}$$

$$= \sum_{\nu=0}^{\infty} v_\nu z^{\nu} + \sum_{\nu=0}^{\infty} u_\nu z^{-\nu}$$

$$= v(z) + u(z^{-1}) . \quad \square$$

Taken together, these three lemmas give the previously stated theorem which provides a simple characterization for the pole preserving mappings we would like to use in a preprocessing scheme. The discussion of section 2 can now

be extended at once to the case where the tail sum operator (38) is replaced by any operator $\Gamma = \gamma(E)$ for which $\gamma(z)$ is analytic and zero free on the unit disc.

4. NOISE SUPPRESSION

Let $a \neq 0$ and z be given with $|z| < 1$, and let

$$(51) \quad w_k = az^k + \epsilon_k, \quad k=0,1,2,\ldots$$

where $\epsilon_0, \epsilon_1, \epsilon_2, \ldots$ are independent random variables with common mean

$$< \epsilon_k > = 0, \quad k=0,1,2,\ldots$$

and common variance

$$< \epsilon_k^2 > = \sigma^2, \quad k=0,1,2,\ldots \quad .$$

Let

$$\Gamma = \gamma(E) = \gamma_0 + \gamma_1 E + \gamma_2 E^2 + \ldots$$

be a preprocessing operator for which

$$(52) \quad \gamma^2 = |\gamma_0|^2 + |\gamma_1|^2 + |\gamma_2|^2 + \ldots$$

is finite, and let $\mathbf{v} = \Gamma \mathbf{w}$ so that

$$(53) \quad v_k = a\gamma(z)z^k + \delta_k, \quad k=0,1,2,\ldots$$

where

$$(54) \quad \delta_k = \gamma(E)\epsilon_k = \gamma_0\epsilon_k + \gamma_1\epsilon_{k+1} + \gamma_2\epsilon_{k+2} + \ldots, \quad k=0,1,2,$$

is a random variable with mean

$$(55) \quad <\delta_k> = < \sum_{\ell=0}^{\infty} \gamma_\ell \epsilon_{k+\ell} > = \sum_{\ell=0}^{\infty} \gamma_\ell <\epsilon_{k+\ell}> = 0, \quad k=0,1,2,\ldots$$

and variance

$$(56)\quad \langle|\delta_k|^2\rangle=\Big\langle \sum_{\ell,\mu=0}^{\infty} \bar{\gamma}_\ell \gamma_\mu \epsilon_{k+\ell}\epsilon_{k+\mu}\Big\rangle= \sum_{\ell,\mu=0}^{\infty} \bar{\gamma}_\ell \gamma_\mu \langle \epsilon_{k+\ell}\epsilon_{k+\mu}\rangle$$

$$= \sum_{\ell=0}^{\infty} |\gamma_\ell|^2\, \sigma^2=\gamma^2\sigma^2,\quad k=0,1,2,\ldots \quad .$$

The processed noise δ_0, δ_1, δ_2,... is correlated with

$$(57)\quad \langle\bar{\delta}_k\delta_{k+p}\rangle= \sum_{\ell,\mu=0}^{\infty} \bar{\gamma}_\ell \gamma_\mu \epsilon_{k+\ell}\ \epsilon_{k+p+\mu}$$

$$=(\bar{\gamma}_p\gamma_0+\bar{\gamma}_{p+1}\gamma_1 + \bar{\gamma}_{p+2}\gamma_2+\ldots)\sigma^2,\quad k,p=0,1,2,\ldots \quad .$$

We would like to develop some quantitative measure of the tendency of the preprocessing scheme to suppress the noise. We shall use $|az^k|^2$ as a measure of the signal present in the k-th component of $\mathbf{w}$ with $\langle\epsilon_k^2\rangle = \sigma^2$ being a corresponding measure of the noise. Analogously, $|a\gamma(z)z^k|^2$ gives a measure of the signal in the k-th component of $\mathbf{v} = \Gamma\mathbf{w}$ with $\langle|\delta_k|^2\rangle = \gamma^2\sigma^2$ being the corresponding measure of the noise. The preprocessing scheme Γ thus improves the k-th component signal-to-noise ratio by the factor

$$\frac{|a\gamma(z)z^k|^2/(\gamma^2\sigma^2)}{|az^k|^2/\sigma^2} = |\gamma(z)|^2/\gamma^2$$

Since this ratio is independent of k, we see that

$$(58)\quad g(z) = |\gamma(z)|^2/\gamma^2 = \frac{|\gamma_0+\gamma_1 z+\gamma_2 z^2+ \ldots |^2}{|\gamma_0|^2+|\gamma_1|^2+|\gamma_2|^2+\ldots}$$

provides us with a quantitative measure of the tendancy of Γ to enhance the signal-to-noise ratio for a power sequence based on the pole z. When $g(z) > 1$, the noise is suppressed, and we refer to the set

$$(59)\quad \mathcal{A}_\gamma = \{z \in \mathbf{C} : |z| < 1 \text{ and } g(z) > 1\}$$

as the region of pole amplification associated with $\Gamma = \gamma(E)$.

Suppose now that

$$(60)\quad w_k = \sum_{\nu=1}^{n} a_\nu z_\nu^k + \epsilon_k,\ k=0,1,2,\ldots$$

where $a_\nu \neq 0$ and $|z_\nu| < 1$ for $\nu=1,\ldots, n$, where $z_1,\ldots, z_n$ are distinct, and where $\epsilon_0,\epsilon_1,\epsilon_2,\ldots$ are again independent random variables with zero mean and common variance σ^2. If we compute

$$(61)\quad v_k = \gamma(E)w_k = \sum_{\nu=1}^{n} a_\nu \gamma(z_\nu) z_\nu^k + \delta_k,\ k=0,1,2,\ldots$$

(with $\delta_0,\delta_1,\delta_2,\ldots$ again given by (54)) we will expect to enhance the pole z_ν relative to the noise provided that z_ν lies within the region of amplification $\mathcal{A}_\gamma$ with the degree of enhancement depending on the size of $g(z_\nu)$. In practice, the poles $z_1,\ldots, z_n$ are unknown, but we often

have some a priori knowledge of the portion of the unit disc in which they are most likely to be found. In such situations we might reasonably select a preprocessing scheme which is specifically designed for signals of the type we expect to process, i.e., we choose $\Gamma = \gamma(E)$ so as to make $g(z)$ as large as possible in the region of the unit circle where we expect to find $z_1, \ldots, z_n$. An upper bound on the possible size of $g(z)$ is provided by

THEOREM 2. Let the operator $\Gamma = \gamma(E)$ satisfy (32) - (34), (37) with (52) being finite and let $g(z)$ be the corresponding signal-to-noise amplification factor (58). Then

$$(62) \quad g(z_0) \leq 1/(1-|z_0|^2), \qquad |z_0| < 1$$

with equality holding if and only if

$$(63) \quad \gamma(z) = a/(1-\bar{z}_0 z), \quad |z| < 1$$

where a is a nonzero constant so that $\gamma(z)$ has a simple pole at the point of inversion $z = 1/\bar{z}_0$ of z_0 relative to the unit circle and

$$(64) \quad \Gamma y_k = a(y_k + \bar{z}_0 \cdot y_{k+1} + \bar{z}_0^2 \cdot y_{k+2} + \ldots), \quad k=0,1,2,\ldots .$$

Proof. Using Cauchy's inequality we find

$$(65)\quad |\gamma(z)|^2 = \Big|\sum_{k=0}^{\infty} \gamma_k\, z^k\Big|^2$$

$$\leq \sum_{k=0}^{\infty} |\gamma_k|^2 \sum_{k=0}^{\infty} |z^k|^2$$

$$= \gamma^2/(1-|z|^2)$$

with equality holding if and only if $\boldsymbol{\gamma}$ is a scalar multiple of the power sequence $(1,\overline{z},\overline{z}^2,\dots)$. In conjunction with (58) this gives (62) - (64). □

The bound (62) provides a natural limit to the signal-to-noise amplification we can achieve through the use of a preprocessing scheme. When the sampling interval T is chosen so large that the pole z_ν lies deep within the unit circle, i.e., $z_\nu \approx 0$, the successive powers $z_\nu^2, z_\nu^3, \dots$ decay so rapidly that substantial noise suppression is impossible and we find

$$g(z_\nu) \leq 1/(1-|z_\nu|^2) \approx 1 \quad .$$

On the other hand, if T is so small that z_ν lies near the rim of the unit circle, i.e., $|z_\nu| < 1$ but $|z_\nu| \approx 1$, then the maximum of

$$g(z_\nu) = 1/(1-|z_\nu|^2) \approx 1/[2(1-|z_\nu|)]$$

will be large and a suitable preprocessing scheme can achieve a substantial suppression of the noise, cf. Fig.1.

Suppose now that we have some a priori knowledge that the poles $z_1, \ldots, z_n$ which we seek all lie in some subset $\mathcal{A}$ of the unit disc, e.g., if the sampling interval T is small we expect to find $z_1, \ldots, z_n$ in a tight cluster near z=1 so we might take $\mathcal{A}$ to be the lens shaped set

$$\mathcal{A} = \{z \in \mathbf{C} : |z| < 1 \text{ and } |z-1| < \rho\} \tag{66}$$

obtained by intersecting the unit disc with the disc of radius ρ which is centered at z=1. Once $\mathcal{A}$ is chosen, we would like to determine a corresponding preprocessing operator $\Gamma = \gamma(E)$ which is optimal in the sense that the minimum value taken by g(z) as z ranges over $\mathcal{A}$ (i.e., the smallest signal-to-noise amplification at any pole location we might possibly encounter) is as large as possible. When $\mathcal{A} = \{z_0\}$, the optimal operator Γ is given by (63) - (64). In more realistic situations such as (66), there is no known procedure for constructing an optimal Γ. Nevertheless, we can develop good if not best preprocessors for certain natural choices of $\mathcal{A}$ and we shall now proceed to show how this is done.

Suppose first that $\mathcal{A}$ is given by (66) with ρ being small. When $\gamma(z)$ is given by (39) the contour lines of $|\gamma(z)|^2$ are circles centered at z=1 with $|\gamma(z)|^2 > \rho^{-2}$ when $0 < |z-1| < \rho$. We might thus expect the tail sum operator $S = \gamma(E)$ of (38) to be an ideal choice. Unfortunately, the corresponding

γ^2 of (52) is not finite when $\gamma(z)$ is given by (39), so we are forced to consider related approximations.

One very natural approximation is obtained by simply truncating the infinite series (39) so as to obtain (42) with N being a positive integer. The corresponding preprocessing scheme $\mathbf{u} = \gamma(E)\mathbf{y}$ is then the local averaging procedure of (40), and the signal-to-noise enhancement factor is given by

$$(67) \quad g(z) = |1+z+ \ldots +z^{N-1}|^2/N = |1-z^N|^2/(|1-z|^2N) .$$

By using Cauchy's inequality we see that this operator is optimal in the sense that it maximizes

$$g(1) = \frac{|\gamma_0+\gamma_1 z+ \ldots +\gamma_{N-1}z^{N-1}|^2}{|\gamma_0|^2+|\gamma_1|^2+ \ldots +|\gamma_{N-1}|^2}$$

as $\gamma(z)$ ranges over the set of all (nonzero) polynomials having degree less than N. If we let

$$\zeta_k = \exp(i \cdot 2\pi k/N), \; k=1,2,\ldots, N-1, \; (i^2 = -1)$$

denote the N-th roots of unity other than z=1, we may write (67) in the form

$$g(z) = |z-\zeta_1|^2 \ldots |z-\zeta_{N-1}|^2/N$$

and thereby conclude that the contour lines of $g(z)$ coincide with the equipotentials of the two dimensional field which

results when unit charges are placed at $z = \zeta_1, \zeta_2, \ldots, \zeta_{N-1}$ in the plane. This helps us visualize the region of amplification when N is small, cf. Fig. 2. For large N we neglect z^N in (67) to see that the region of amplification is essentially the lens shaped set (66) which results when we take $\rho = N^{-1/2}$ and that $g(z)$ increases from approximately 1 to its maximum value N as the distance from z to 1 decreases from $N^{-1/2}$ to 0. In this way we see that a large signal-to-noise amplification is possible only when N is large in which case the corresponding region of pole amplification is small. Thus N should be chosen in conjunction with the sampling interval T, i.e., when T is small the z_ν's are tightly clustered near z=1 and a large value of N is appropriate, cf. Fig. 1 and Fig. 3 .

A second natural approximation to (39) is obtained by slightly shifting the pole z=1 of γ to the nearby point $z = R > 1$ so as to make

$$(68) \quad \gamma(z) = 1 + z/R + z^2/R^2 + \ldots = (1-z/R)^{-1} \quad .$$

In this case $\mathbf{u} = \gamma(E)\mathbf{y}$ is given by

$$(69) \quad u_k = y_k + y_{k+1}/R + y_{k+2}/R^2 + \ldots$$

and

$$(70) \quad g(z) = (R^2-1)/|z-R|^2 \quad .$$

The contour lines of $g(z)$ are circles centered at z=R

with $g(z) = K$ on the circle where

$$|z-R| = [(R^2-1)/K]^{1/2} .$$

In particular, the maximum signal-to-noise enhancement is

$$g(1) = (R+1)/(R-1) \approx 2/(R-1)$$

and the region of amplification is that portion of the unit disc $|z| < 1$ which lies within the circle

$$|z-R| = (R^2-1)^{1/2} .$$

Again we observe a fundamental trade off between the size of the region of amplification and the maximum possible g, cf. Fig. 4.

5. CONSTRUCTING NOISE SUPPRESSING PREPROCESSING OPERATORS

In many cases of practical importance we encounter a system which exhibits various modes of oscillation with frequencies which are more or less integral multiples of some fundamental frequency and with all of these modes being similarly damped (cf. Appendix 1), i.e., the s_ν's of (1) are more or less uniformly spaced along some line $\mathrm{Re}\, s = -\alpha$ in the left half plane and the corresponding z_ν's of (4) are more or less uniformly spaced around the circle $|z| = \exp(-\alpha T) < 1$. To aid us in the numerical extraction of the first few z_ν's we would like to use a preprocessing scheme which is optimal on some corresponding sector

$$(71)\quad \mathcal{A} = \{z=re^{i\theta} : \rho < r < 1,\ -\theta < \theta < \theta\}$$

of some annulus of the unit circle with the parameters ρ and θ ($0 < \rho < 1$ and $0 < \theta < \pi$) being determined by the damping coefficient α, the fundamental frequency, the sampling interval, and the number of poles we are trying to find. The preprocessing schemes of (40) and (69) (which are designed for the region (66)) tend to enhance the lowest modes but damp out any others. We must thus devise a preprocessing scheme $\Gamma = \gamma(E)$ for a region of the form(71). We now describe one way in which this can be done.

We first observe that the function

$$(72)\quad \gamma^*(z) = \frac{R^2}{z^2-2Rz\cos\theta+R^2} = \frac{R^2}{(z-Re^{i\theta})(z-Re^{-i\theta})}$$

is analytic inside the circle $|z| = R$ on which its two conjugate poles $z = R\exp(\pm i\theta)$ are found. For $|z| < R$ and $0 < \theta < \pi$ we may write

$$\gamma^*(z) = \frac{1}{2i\sin\theta}\left\{\frac{e^{i\theta}}{1-(z/R)e^{i\theta}} - \frac{e^{-i\theta}}{1-(z/R)e^{-i\theta}}\right\}$$

$$= \frac{1}{2i\sin\theta}\left\{e^{i\theta}\sum_{k=0}^{\infty}(e^{i\theta}z/R)^k - e^{-i\theta}\sum_{k=0}^{\infty}(e^{-i\theta}z/R)^k\right\},$$

and after a bit of simplification this yields the series representation

$$(73)\quad \gamma^*(z) = 1 + \frac{\sin 2\theta}{\sin\theta}\cdot\frac{z}{R} + \frac{\sin 3\theta}{\sin\theta}\cdot\frac{z^2}{R^2} + \cdots$$

for (72).

To obtain our desired preprocessing operator we shall now select a number p such R_j, θ_j pairs with $R_j > 1$ and $0 < \theta_j < \pi$, and then form the product of the corresponding terms (72), i.e., we take

$$(74)\quad \gamma(z) = \prod_{j=1}^{p}\frac{R_j^2}{z^2-2R_jz\cos\theta_j+R_j^2} = \gamma_0+\gamma_1z+\gamma_2z^2+\cdots\ .$$

From (73) - (74) we see that $\gamma(z)$ has the alternative representation

$$(75)\quad \gamma(z) = \prod_{j=1}^{p} \left\{ 1 + \frac{\sin 2\theta_j}{\sin \theta_j} \frac{z}{R_j} + \frac{\sin 3\theta_j}{\sin \theta_j} \cdot \frac{z^2}{R_j^2} + \ldots \right\} ,$$

so that if we define

$$\gamma^{(1)}(z) = 1 + \frac{\sin 2\theta_1}{\sin \theta_1} \frac{z}{R_1} + \frac{\sin 3\theta_1}{\sin \theta_1} \cdot \frac{z^2}{R_1^2} + \ldots$$

and successively compute

$$\gamma^{(\ell)}(z) = \gamma^{(\ell-1)}(z) \cdot \left\{ 1 + \frac{\sin 2\theta_\ell}{\sin \theta_\ell} \cdot \frac{z}{R_\ell} + \frac{\sin 3\theta_\ell}{\sin \theta_\ell} \cdot \frac{z^2}{R_\ell^2} + \ldots \right\} ,$$

for $\ell = 2, 3, \ldots, p$ the resulting $\gamma^{(p)}(z)$ will coincide with (75) and thus (74). We may thus numerically obtain the coefficients $\gamma_0, \gamma_1, \gamma_2, \ldots$ of (74) by using the recursion scheme

$$\gamma_k^{(1)} = \frac{\sin(k+1)\theta_1}{\sin \theta_1} \cdot \frac{1}{R_1^k} \quad , \quad k = 0, 1, 2, \ldots$$

(76)

$$\gamma_k^{(\ell)} = \sum_{j=0}^{k} \gamma_{k-j}^{(\ell-1)} \cdot \frac{\sin(j+1)\theta_\ell}{\sin \theta_\ell} \cdot \frac{1}{R_\ell^j} \quad , \quad \begin{matrix} k = 0, 1, 2, \ldots, \\ \ell = 2, 3, \ldots, p \end{matrix}$$

with

$$(77)\quad \gamma_k = \gamma_k^{(p)}, \quad k = 0, 1, 2, \ldots \quad .$$

Once $\gamma_0, \gamma_1, \gamma_2, \ldots$ are known, we use the corresponding operator

$$\Gamma = \gamma_0 + \gamma_1 E + \gamma_2 E^2 + \ldots$$

for our preprocessing scheme, i.e., we take

$$u_k = \gamma_0 y_k + \gamma_1 y_{k+1} + \gamma_2 y_{k+2} + \ldots \quad .$$

From (74) we see that $\gamma(z)$ will be analytic on the disc $|z| < R$ when

$$R = \min \{R_1, \ldots, R_p\} > 1 \ .$$

By using the Cauchy-Hadamard formula (36) we then see that there are constants $C > 0$, $0 < q < 1$ such that the coefficients $\gamma_0, \gamma_1, \gamma_2, \ldots$ we compute in (79)-(80) will satisfy the bound

$$|\gamma_k| < C \cdot q^k, \quad k=0,1,2,\ldots \ .$$

We thereby infer that the sum

$$\sum_{k=0}^{\infty} |\gamma_k|^2 < \sum_{k=0}^{\infty} |Cq^k|^2 = C^2/(1-q^2)$$

is finite so that the corresponding $g(z)$ of (58) is well defined. Our goal is to determine the pole location parameters R_j, θ_j, $j=1,\ldots, p$ which result in a $g(z)$ which is uniformly large on all of $\mathcal{A}$ in the sense that

$$\min \{ g(z) : z \in \mathcal{A} \}$$

is as large as possible. Since $g(z)$ is the modulus of an

analytic function which is zero free on $\mathcal{A}$, this is equivalent to maximizing the minimum value taken by $g(z)$ as z ranges over the boundary of $\mathcal{A}$.

Using a digital computer in an interactive mode (where it is possible to perturb the pole location parameters R_j, θ_j and immediately see the effect on the contour lines of $g(z)$ we have been able to design good if not best preprocessing operators $\Gamma = \gamma(E)$ for the region (71), cf. Fig. 5. More work is needed in order to systematize this procedure, however.

6. ALTERNATIVE METHODS FOR COMPUTING THE NULL VECTOR

At the heart of transient analysis lies the problem of extracting a suitable null vector $\mathbf{c} = (c_0, c_1, \ldots, c_n)$ for the data matrix $\mathcal{Y}$ of (11). As we have seen, $\mathbf{c}$ can be obtained alternatively as a null vector of certain related matrices such as (25), (26) or, more generally, as a null vector of the corresponding matrices which result when the tail sum operator S is replaced by some other preprocessing operator Γ chosen for its noise suppressing properties. We shall now describe a general procedure of this type in the hopes of obtaining a more robust scheme for computing $\mathbf{c}$ and thereby for computing $z_1, \ldots, z_n$.

Indeed, suppose that we are given the transient sequence $\mathbf{y} = (y_0, y_1, y_2, \ldots)$ and the preprocessing operators $\Gamma_i = \gamma_i(E)$, i=0,1,..., m. We shall numerically generate the sequences $\Gamma_i \mathbf{y}$, i=0,1,..., m and then assemble the matrix

$$(78)\quad \mathcal{Y} = \begin{bmatrix} \Gamma_0 y_0 & \Gamma_0 y_1 & \cdots & \Gamma_0 y_n \\ \Gamma_1 y_0 & \Gamma_1 y_2 & \cdots & \Gamma_1 y_n \\ \vdots & & & \\ \Gamma_m y_0 & \Gamma_m y_1 & \cdots & \Gamma_m y_n \end{bmatrix}$$

Using a singular value decomposition of $\mathcal{Y}$ or an eigenvalue decomposition of $\mathcal{Y}^T\mathcal{Y}$ as described in section 1 we shall compute a null vector $\mathbf{c} = (c_0, c_1, \ldots, c_n)$ for $\mathcal{Y}$ and then

take $z_1, \ldots, z_n$ to be the roots which result from the corresponding factorization (6). Of course, this scheme will be successful only if the operators are chosen in such a manner that $\mathcal{H}$ has the same null space as the data matrix $\mathcal{Y}$. This imposes certain constraints on the operators $\Gamma_0, \Gamma_1, \ldots, \Gamma_m$ which we shall now explore in some detail.

Let $\gamma_0(z), \gamma_1(z), \ldots, \gamma_m(z)$ be analytic on the unit disc $|z| < 1$. We shall say that these functions satisfy the Haar condition provided that no linear combination

$$(79) \quad \gamma_*(z) = b_0\gamma_0(z) + b_1\gamma_1(z) + \ldots + b_m\gamma_m(z)$$

of these functions has more than m zeros in $|z| < 1$ unless $b_0, b_1, \ldots, b_m$ all vanish. After illustrating this concept with several examples we shall show that this is exactly the property which is needed in order to insure that $\mathcal{H}$ and $\mathcal{Y}$ share the same null space.

Example 1. The functions

$$\gamma_i(z) = z^i, \quad i=0,1,\ldots, m$$

satisfy the Haar condition on the unit disc since (79) reduces to the polynomial

$$\gamma_*(z) = b_0 + b_1 z + \ldots + b_m z^m$$

which has at most m zeros unless $b_0 = b_1 = \ldots = b_m = 0$. In this case

$$\Gamma_i y_j = E^i y_j = y_{i+j}$$

so that $\mathcal{Y}$ is the matrix which results when all but the first m+1 rows of $\mathcal{Y}$ are deleted.

Example 2. When $\gamma(z)$ is analytic and zero free on the unit disc the functions

$$\gamma_i(z) = z^i \gamma(z) , \quad i=0,1,\ldots, m \tag{80}$$

satisfy the Haar condition there since (79) reduces to the product

$$\gamma_*(z) = \gamma(z) \cdot [b_0 + b_1 z + \ldots + b_m z^m]$$

which has at most m zeros unless $b_0=b_1= \ldots = b_m = 0$. In this case

$$\Gamma_i y_j = E^i \gamma(E) y_j = \gamma(E) y_{i+j}$$

so that $\mathcal{Y}$ is obtained by generalizing (25) to the case where S^p is replaced by $\gamma(E)$.

Example 3. Again let $\gamma(z)$ be analytic on the unit disc and assume that $\gamma(z)$ is also one-to-one, i.e., that $\gamma(z) = \gamma(z')$ only when $z = z'$. If

$$\gamma_i(z) = \gamma(z)^i , \quad i=0,1,\ldots, m \tag{81}$$

then (79) reduces to the polynomial

$$\gamma_*(z) = b_0 + b_1 \gamma(z) + b_2 \gamma(z)^2 + \ldots + b_m \gamma(z)^m$$

of degree m in $\gamma(z)$. Since $\gamma(z)$ is one-to-one, $\gamma_*(z)$ will have at most m zeros unless $b_0 = b_1 = \ldots = b_m = 0$, and so the Haar condition is satisfied once again. In this case

$$\Gamma_i y_j = \gamma(E)^i y_j$$

so that $\mathcal{Y}$ is obtained by generalizing (26) to the case where S is replaced by $\gamma(E)$.

Example 4. Let $\zeta_0, \zeta_1, \ldots, \zeta_m$ be distinct complex numbers with $|\zeta_i| < 1$ for each i, and let

$$(82) \quad \gamma_i(z) = 1/(1-\overline{\zeta}_i z) , \quad i=0,1,\ldots, m .$$

We may express the corresponding expression (79) in the form

$$\gamma_*(z) = b_0/(1-\overline{\zeta}_0 z) + \ldots + b_m/(1-\overline{\zeta}_m z)$$

$$= \frac{d_0 + d_1 z + \ldots + d_m z^m}{(1-\overline{\zeta}_0 z)(1-\overline{\zeta}_1 z)\ldots(1-\overline{\zeta}_m z)}$$

with $d_0, d_1, \ldots, d_m$ depending linearly on $b_0, b_1, \ldots, b_m$. Clearly $\gamma_*(z)$ has at most m zeros unless $d_0=d_1= \ldots = d_m = 0$ in which case $b_0=b_1= \ldots =b_m=0$ also, so the $\gamma_i(z)$'s satisfy the Haar condition on the unit disc. Using Theorem 2 we see that the operator

$$\Gamma_i = \gamma_i(E) = I + \overline{\zeta}_i E + \overline{\zeta}_i{}^2 E^2 + \ldots$$

provides the maximum possible improvement in the signal-to-noise ratio for a pole at $z=\zeta_i$, and thus such a pole will be strongly represented in the i-th row of $\mathcal{Y}$, i=0,1,..., m.

Using this $\mathcal{Y}$ to determine the null vector **c** may be preferable to using (25), (26), or the generalization of Examples 2, 3 above when the system poles are well separated and it is difficult or impossible to find a single operator Γ which has a region of amplification containing all of these poles.

Example 5. Again let $\zeta_0, \zeta_1, \ldots, \zeta_m$ be distinct complex numbers with $|\zeta_i| < 1$ for each i and let

$$\gamma_i(z) = \prod_{j \neq i} \frac{z - \zeta_j}{\zeta_i - \zeta_j} , \quad i=0,1,\ldots, m$$

be the Lagrange interpolating polynomials of degree m which are based on the m+1 points $\zeta_0, \zeta_1, \ldots, \zeta_m$ so that

$$\gamma_i(\zeta_j) = \begin{cases} 1 & \text{if} \quad i=j = 0,1,\ldots, m \\ 0 & \text{if} \quad i \neq j . \end{cases}$$

Since $\gamma_0(z), \gamma_1(z), \ldots, \gamma_m(z)$ are linearly independent polynomials of degree m, the Haar condition is satisfied. The application of the operator $\Gamma_i = \gamma_i(E)$ to a given sequence $\mathbf{y} = (y_0, y_1, \ldots)$ can be effected by successively applying the m operators $E - \zeta_j I$, $j \neq i$, and then suitably scaling the resulting sequence.

We note that in the special case where **y** has the representation (3), where m=n-1, and where we take $\zeta_0 = z_1$, $\zeta_1 = z_2, \ldots, \zeta_m = z_n$, we find that the elements of $\mathcal{Y}$ are given by

$$\Gamma_i y_j = \sum_{\nu=1}^{n} a_\nu \gamma_i(z_\nu) z_\nu^{\,j} = a_{i+1}\, z_{i+1}^{\,j} \quad .$$

Thus $\mathcal{Y}$ is a row scaled version of the Vandermonde matrix based on $z_1, \ldots, z_n$.

THEOREM 3. Let $\Gamma_i = \gamma_i(E)$ where $\gamma_i(z)$ is analytic for $|z| < 1$, $i=0,1,\ldots, m$ and assume that $\gamma_0(z), \gamma_1(z), \ldots, \gamma_m(z)$ satisfy the Haar condition on the unit disc. Let $\mathbf{y} = (y_0, y_1, y_2, \ldots)$ be a transient sequence having the representation (3) with $a_\nu \neq 0$ and $|z_\nu| < 1$ for each $\nu=1,\ldots, n$, with $z_1, \ldots, z_n$ being distinct, and with $n \leq m+1$. Then $\mathbf{c} = (c_0, c_1, \ldots, c_n)^T$ is a null vector of the matrix $\mathcal{Y}$ of (78) if and only if $\mathbf{c}$ is also a null vector of the data matrix $\mathcal{Y}$ of (11).

Proof. Using (3) we see that for any choice of $c_0, c_1, \ldots, c_n$ we have

$$(83)\quad c_0\mathcal{Y}_{i0}+c_1\mathcal{Y}_{i1}+\ldots+c_n\mathcal{Y}_{in}=(c_0+c_1E+\ldots+c_nE^n)\gamma_i(E)\sum_{\nu=1}^{n}a_\nu z_\nu^{\,k}\Big|_{k=0}$$

$$= \sum_{\nu=1}^{n} a_\nu (c_0+c_1 z_\nu+\ldots+c_n z_\nu^{\,n})\gamma_i(z_\nu).$$

Now if $\mathbf{c}$ is a null vector of $\mathcal{Y}$, the factorization (6) must hold so

$$c_0\mathcal{Y}_{i0}+c_1\mathcal{Y}_{i1}+\ldots+c_n\mathcal{Y}_{in} = 0, \quad i=0,1,\ldots, m$$

and thus $\mathbf{c}$ is also a null vector of $\mathcal{Y}$. Conversely, if $\mathbf{c}$

is a null vector of $\mathcal{Y}$ the left side of (83) vanishes for each $i=0,1,\ldots, m$ and therefore

$$(84)\quad \sum_{\nu=1}^{n} a_\nu(c_0+c_1z_\nu+\ldots+c_nz_\nu^n)[b_0\gamma_0(z_\nu)+b_1\gamma_1(z_\nu)+\ldots+b_m\gamma_m(z_\nu)]=0$$

for every choice of $b_0, b_1,\ldots, b_m$. Since $\gamma_0(z),\gamma_1(z),\ldots,\gamma_m(z)$ satisfy the Haar condition, we can choose the b_i's so as to make (79) interpolate any given $m+1$ points, and in particular (since $m \geq n-1$) we can choose these coefficients so as to make

$$(85)\quad b_0\gamma_0(z_\nu)+b_1\gamma_1(z_\nu)+\ldots+b_m\gamma_m(z_\nu) = \overline{a_\nu(c_0+c_1z_\nu+\ldots+c_nz_\nu^n)},\ \nu=1,\ldots, n\ .$$

Upon substituting (85) into (84) we find

$$\sum_{\nu=1}^{n} |a_\nu(c_0+c_1z_\nu+\ldots+c_nz_\nu^n)|^2 = 0$$

so that

$$c_0y_i+c_1y_{i+1}+\ldots+c_ny_{i+n} = \sum_{\nu=1}^{n} a_\nu(c_0+c_1z_\nu+\ldots+c_nz_\nu^n)z_\nu^i=0, i=0,1,\ldots\ ,$$

i.e., $\mathbf{c}$ is a null vector of $\mathcal{Y}$. □

The ideas leading to the construction of (28) lead to the somewhat more general result which contains Theorem 3 as a special case.

THEOREM 4. Let $\Gamma_i = \gamma_i(E)$ where $\gamma_i(z)$ is analytic for $|z| < 1$, $i=0,1,\ldots, m$ and assume that $\gamma_0(z)$, $\gamma_1(z)$, ..., $\gamma_m(z)$ satisfy the Haar condition on the unit disc. Let $\Delta_j = \delta_j(E)$ where $\delta_j(z)$ is analytic for $|z| < 1$, $j=0,1,\ldots n$ and assume that $\delta_0(z)$, $\delta_1(z)$, ..., $\delta_n(z)$ also satisfy the Haar condition on the unit disc. Let $\mathbf{y}$ be a transient sequence having the representation (3) with $a_\nu \neq 0$ and $|z_\nu| < 1$ for $\nu=1,\ldots,n$, with $z_1,\ldots, z_n$ being distinct, and with $n \leq m+1$. Then $\mathbf{d}=(d_0,d_1,\ldots, d_n)^T$ is a null vector of the $(m+1) \times (n+1)$ matrix

$$(86)\quad \mathcal{M} = \begin{bmatrix} \Gamma_0\Delta_0 y_0 & \Gamma_0\Delta_1 y_0 \cdots \Gamma_0\Delta_n y_0 \\ \Gamma_1\Delta_0 y_0 & \Gamma_1\Delta_1 y_0 \cdots \Gamma_1\Delta_n y_0 \\ \vdots & \\ \Gamma_m\Delta_0 y_0 & \Gamma_m\Delta_1 y_0 \cdots \Gamma_m\Delta_n y_0 \end{bmatrix}$$

if and only if the corresponding characteristic function

$$(87)\quad q(z) = d_0\,\delta_0(z) + d_1\delta_1(z) + \ldots + d_n\delta_n(z)$$

vanishes at $z_1,\ldots, z_n$ and is elsewhere nonzero for $|z| < 1$.

Proof. Using the representation (3) we see that

$$(88)\quad \mathbf{b}^T\mathcal{M}\mathbf{d} = \sum_{\nu=1}^{n} a_\nu [b_0\gamma_0(z_\nu)+b_1\gamma_1(z_\nu)+\ldots+b_m\gamma_m(z_\nu)] \\ [d_0\delta_0(z_\nu)+d_1\delta_1(z_\nu)+\ldots+d_n\delta_n(z_\nu)]$$

for every choice of $\mathbf{b}^T=(b_0,b_1,\ldots, b_m)$ and $\mathbf{d}^T=(d_0,d_1,\ldots,d_n)$. If $\mathbf{d}$ is chosen so that (87) has $z_1,\ldots, z_n$ as roots, then the left side of (88) must vanish for every choice of $\mathbf{b}$ (and in particular for the choice $\mathbf{b} = \mathfrak{M}\,\mathbf{d}$), and we conclude that $\mathbf{d}$ is a null vector of $\mathfrak{M}$.

Conversely, suppose that $\mathbf{d}\neq 0$ is a null vector of $\mathfrak{M}$ so that the right hand side of (88) vanishes for each choice of $\mathbf{b}$. Since $\gamma_0(z)$, $\gamma_1(z),\ldots, \gamma_m(z)$ satisfy the Haar condition and $n \leq m+1$ it is possible to choose $\mathbf{b}$ so that

$$b_0\gamma_0(z_\nu)+b_1\gamma_1(z_\nu)+\ldots+b_m\gamma_m(z_\nu) = \overline{a_\nu[d_0\delta_0(z_\nu)+d_1\delta_1(z_\nu)+\ldots+d_n\delta_n(z_\nu)]}$$

when $\nu=1,\ldots, n$. It then follows that

$$\sum_{\nu=1}^{n}\left|a_\nu[d_0\delta_0(z_\nu)+d_1\delta_1(z_\nu)+\ldots+d_n\delta_n(z_\nu)]\right|^2 = 0$$

and thus (87) must vanish at $z_1,\ldots, z_n$. Moreover, since $\delta_0(z)$, $\delta(z),\ldots, \delta_n(z)$ also satisfy the Haar condition, the characteristic function (87) has no other zeros in the unit disc. □

Example 6. Let

$$\gamma_i(z) = z^i,\ i=0,1,\ldots, m$$

and let

$$\delta_j(z) = \gamma(z)^j,\ j = 0,1,\ldots,\ n$$

where $\gamma(z)$ is analytic and one-to-one for $|z| < 1$. As noted in Examples 1,3 above the $\gamma_i(z)$'s and $\delta_j(z)$'s satisfy the Haar condition on the unit disc. If we set $\Gamma = \gamma(E)$ we may write

$$\Gamma_i = \gamma_i(E) = E^i,\ i=0,1,\ldots,\ m$$

$$\Delta_j = \delta_j(E) = \Gamma^j,\ j=0,1,\ldots,\ n$$

so that

$$(89)\quad \mathfrak{M} = \begin{bmatrix} y_0 & \Gamma y_0 & \ldots & \Gamma^n y_0 \\ y_1 & \Gamma y_1 & \ldots & \Gamma^n y_1 \\ \vdots & \vdots & & \vdots \\ y_m & \Gamma y_m & \ldots & \Gamma^n y_m \end{bmatrix}$$

We then conclude that $\mathbf{d} = (d_0, d_1,\ldots, d_n)$ is a null vector of $\mathfrak{M}$ if and only if $\gamma(z_1),\ldots, \gamma(z_n)$ are the roots of the polynomial $d_0 + d_1 w +\ldots+ d_n w^n$, and since γ is one-to-one the z_ν's are uniquely determined from $\mathbf{d}$. We observe that (28) is the specialization of (89) to the case where $\gamma(z) = 1/(1-z)$ as used in Jain's analysis.

Example 7. Let

$$\gamma_i(z) = z^i,\ i=0,1,\ldots,\ m$$

and let $\delta_0(z)$, $\delta_1(z),\ldots,$ $\delta_n(z)$ be the Chebyshev polynomials which may be defined and computed by means of the three term recursion

$$
\begin{aligned}
&\delta_0(z) = 1 \\
&\delta_1(z) = z \\
&\delta_j(z) = 2z\delta_{j-1}(z) - \delta_{j-2}(z), \quad j=2,3,\ldots, n\ .
\end{aligned}
\tag{90}
$$

Since $\delta_j(z)$ is a polynomial of exact degree j, the Haar condition is clearly satisfied. By using (90) we see that it is possible to numerically generate the sequences

$$\Delta_j \mathbf{y} = \delta_j(E)\mathbf{y}\ ,\ j=0,1,\ldots,\ n$$

when $\mathbf{y}$ is given by successively computing

$$
\begin{aligned}
&\Delta_0 y_i = y_i,\ i=0,1,2,\ldots \\
&\Delta_1 y_i = y_{i+1},\ i=0,1,2,\ldots \\
&\Delta_j y_i = 2\Delta_{j-1} y_{i+1} - \Delta_{j-2} y_i,\ i=0,1,2,\ldots,\ j=2,3,\ldots,\ n.
\end{aligned}
$$

When $\mathbf{y}$ has the representation (3) and $m \geq n-1$, $z_1,\ldots,z_n$ will be the roots of the characteristic polynomial (87) (which is now parametrized using the Chebyshev polynomials) if and only if $\mathbf{d} = (d_0,d_1,\ldots,\ d_n)^T$ is a null vector of the matrix

$$
\mathcal{M} = \begin{bmatrix}
\Delta_0 y_0 & \Delta_1 y_0 & \cdots & \Delta_n y_0 \\
\Delta_0 y_1 & \Delta_1 y_1 & \cdots & \Delta_n y_1 \\
\vdots & \vdots & & \vdots \\
\Delta_0 y_m & \Delta_1 y_m & & \Delta_n y_m
\end{bmatrix}
$$

7. THE PROBLEM OF EXTRANEOUS ROOTS

If in the process of numerically analyzing a given transient $\mathbf{y} = (y_0, y_1, y_2 \ldots)$ of the form (3) we overestimate the system order n, i.e., we carry out the calculations with n replaced by some larger integer $m=n+r$, we obtain a characteristic polynomial

$$(91)\quad d(z) = d_0 + d_1 z + \ldots + d_m z^m,\ d_m \neq 0$$

of degree $m > n$. The following theorem provides several equivalent characterizations of $d(z)$ and shows that $z_1, \ldots, z_n$ are to be found among its m roots.

THEOREM 5. Let $\mathbf{y} = (y_0, y_1, y_2 \ldots)$ have the representation (3) with $a_1, \ldots, a_n$ being nonzero and with $z_1, \ldots, z_n$ being distinct. Let $c_0, c_1, \ldots, c_n$ be the coefficients of the characteristic polynomial $c(z)$ of (6) having $z_1, \ldots, z_n$ as its roots. Let $\mathbf{d} = (d_0, d_1, \ldots, d_m)^T$ with $m=n+r$ and $r \geq 0$, and let $d(z)$ be the corresponding polynomial (91). Then the following are equivalent:

(i) $\mathbf{d}$ is a (right) null vector for the data matrix

$$(92)\quad \mathcal{Y}_m = \begin{bmatrix} y_0 & y_1 & \cdots & y_m \\ y_1 & y_2 & \cdots & y_{m+1} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \end{bmatrix}$$

(ii) $d(E)\,\mathbf{y} = \mathbf{0}$,

(iii) $d(z_\nu) = 0$ for each $\nu=1,\ldots, n$,

(iv) $d(z)$ has the factorization

$$d(z) = (b_0 + b_1 z + \ldots + b_r z^r) \cdot c(z)$$

for some choice of the coefficients $b_0, b_1, \ldots, b_r$,

(v) $\mathbf{d}^T$ lies in the row space of the $(r+1) \times (m+1)$ matrix

$$(93)\quad C = \begin{bmatrix} c_0 c_1 \cdots c_n & & \\ c_0\ c_1 \cdots c_n & & \\ \ddots & \ddots & \\ & c_0\ c_1 \cdots c_n \end{bmatrix},$$

i.e.,

$$(94)\quad \mathbf{d}^T = \mathbf{b}^T \cdot C$$

for some choice of $\mathbf{b}^T = (b_0, b_1, \ldots, b_r)$.

Proof. From the identity

$$d(E)\mathbf{y} = d_0\mathbf{y} + d_1 E\mathbf{y} + \ldots + d_m E^m \mathbf{y} = \mathcal{Y}_m\, \mathbf{d}$$

we see that (i) and (ii) are equivalent. Next, since the power sequences $(1, z_\nu, z_\nu^2, \ldots)$, $\nu = 1, \ldots, n$ are linearly independent, and since

$$d(E)y_k = d(E) \sum_{\nu=1}^{n} a_\nu z_\nu^k = \sum_{\nu=1}^{n} [a_\nu d(z_\nu)] z_\nu^k,\quad k=0,1,2,\ldots$$

we see that (ii) holds if and only if $a_\nu d(z_\nu) = 0$ for each

$\nu = 1,\ldots, n$ which is equivalent to (iii) and thus to (iv) since the a_ν's are nonzero. Finally, by equating the coefficients of like powers of z in the identity

$$d_0+d_1z+\ldots+d_mz^m = (b_0+b_1z+\ldots+b_rz^r)(c_0+c_1z+\ldots+c_nz^n)$$

$$= b_0c(z) + \quad b_1zc(z)+ \ldots + b_rz^rc(z)$$

we see that (iv) is equivalent to (v). □

NOTE. Since the conditions (iii), (iv), and (v) above do not depend upon any particular choice of $\mathbf{y}$, these equivalent conditions imply that (i) and (ii) must both hold for every possible nondegenerate choice of $\mathbf{y}$ having the form (3) with the same z_ν's.

There are several ways which can be used to numerically generate a family of vectors

$$(95)\quad \mathbf{d}_i^T = (d_{i0}, d_{i1},\ldots, d_{im}),\ i=0,1,\ldots, r$$

of the form characterized in Theorem 5. We have already observed that if $\mathbf{y}$ is a nondegenerate transient sequence of the form (3) and we perform a singular value decomposition of the data matrix $\mathcal{Y}_m$ with $m = n+r$, then $r+1$ of the singular values will be zero and the corresponding right singular vectors will be null vectors of $\mathcal{Y}_m$ and thus serve as the $\mathbf{d}_i$'s . In some cases, it is convenient to observe a number of transient sequences $\mathbf{y}_0, \mathbf{y}_1,\ldots, \mathbf{y}_r$

which result from different initial excitations of the same underlying physical system. We can then numerically generate a null vector $\mathbf{d}_i$ for the data matrix $\mathcal{Y}_m$ of (92) constructed using the i-th transient sequence $\mathbf{y}_i$, i=0,1,..., r. The same procedure can be used to generate the $\mathbf{d}_i$'s when only one nondegenerate transient $\mathbf{y}$ is known provided that we first generate suitable auxiliary sequences $\mathbf{y}_i = \Gamma_i \mathbf{y}$, i=0,1,..., r from $\mathbf{y}$ by applying preprocessing operators $\Gamma_0, \Gamma_1, \ldots, \Gamma_r$. Indeed, in principle, we can map any given nondegenerate transient $\mathbf{y}$ having component (3) into an arbitrary transient $\mathbf{u}$ having components

$$u_k = \sum_{\nu=1}^{n} a_\nu^* \, z_\nu^k \ , \ k=0,1,2,\ldots$$

(for the same underlying system) by using any preprocessing operator $\Gamma = \gamma(E)$ constructed from a function $\gamma(z)$ which is analytic on $|z| < 1$ and which interpolates the points

$$\gamma(z_\nu) = a_\nu^*/a_\nu \ , \ \nu = 1,\ldots, n \ .$$

We would like to have some numerical procedure for obtaining the system poles $z_1,\ldots, z_n$ from such a collection of vectors (95). In principle, we could simply factor each of the polynomials

$$(96)\ d_i(z) = d_{i0} + d_{i1} z + \ldots + d_{im} z^m \ , \ i=0,1,\ldots, r$$

and choose as $z_1, \ldots, z_n$ the roots which are common to all of them (assuming that the $\mathbf{d}_i$'s have been suitably restricted so as to rule out the possibility of an extraneous zero held in common by each of the polynomials (96) .) In practice, however, the roots are subject to slight perturbations due to noise, computer roundoff, etc., and instead of finding n roots which are held in common by the polynomials (96), we obtain n clusters of roots (near $z_1, \ldots, z_n$) which must be suitably averaged. For this reason, we would like to have some way to process the vectors $\mathbf{d}_0, \mathbf{d}_1, \ldots, \mathbf{d}_r$ so as to obtain directly a good estimate of the coefficients $\mathbf{c} = (c_0, c_1, \ldots, c_n)$ of the characteristic polynomial (6) having $z_1, \ldots, z_n$ as roots. One very nice scheme for doing this has been published recently by Henderson [4], and we shall now expand upon his work.

Suppose then that we have been given a collection (95) with $z_1, \ldots, z_n$ being common roots of each of the corresponding polynomials (96). By using Theorem 5-v on a row-by-row basis we see that the matrix

$$(97)\quad D = \begin{bmatrix} d_{00} & d_{01} & \cdots & d_{0m} \\ d_{10} & d_{11} & \cdots & d_{1m} \\ \vdots & & & \\ d_{r0} & d_{r1} & \cdots & d_{rm} \end{bmatrix}$$

(having $\mathbf{d}_0^T, \mathbf{d}_1^T, \ldots, \mathbf{d}_r^T$ as rows) can be factored in tne form

(98) $D = B\,C$

where B is an (r+1)x (r+1)square matrix

$$(99)\quad B = \begin{bmatrix} b_{00} & b_{01} & \cdots & b_{0r} \\ b_{10} & b_{11} & \cdots & b_{1r} \\ \vdots & & & \\ b_{r0} & b_{r1} & \cdots & b_{rr} \end{bmatrix}$$

(having rows which correspond to the polynomial factors of Theorem 5-iv) and where C is the (r+1) x (m+1) matrix (93). Our goal is to extract the parameters $c_0, c_1, \ldots, c_n$ of C directly from the matrix D.

Henderson's scheme for finding the c_i's begins with the use of Gaussian elimination with partial pivoting to systematically zero out the elements of the matrix D having indices i, j with i > j, i.e., those elements which lie below the principle or left diagonal drawn through d_{00}, d_{11}, ... This process replaces D by the matrix LPD where P is obtained by suitably permuting tne rows of tne (r+1) x (r+1) identity matrix and where L is an (r+1) x (r+1) lower triangular matrix having 1's along its diagonal, cf. [8, Chapter 1]. The application of P serves to permute the rows of D and the subsequent application of L then serves to carry out the elementary row operations which introduce the desired zero structure. Henderson's scheme then

continues by using additional row operations to zero out those elements of LPD having indices i, j with $j > i+m-r$, i.e., those elements which lie above the right diagonal drawn up through d_{rm}, $d_{r-1,m-1}$, ... In this way the matrix LPD is replaced by the matrix ULPD where U is an $(r+1) \times (r+1)$ upper triangular matrix having 1's along its diagonal. The remaining matrix is then a row scaled version of C, i.e.,

$$\text{(100)} \quad \text{ULPD} = \text{SC}$$

where

$$\text{(101)} \quad S = \begin{bmatrix} s_0 & & & O \\ & s_1 & & \\ & & \ddots & \\ O & & & s_r \end{bmatrix}$$

The i-th row of ULPD thus contains $s_i c_0$, $s_i c_1$,..., $s_i c_n$ in columns $j = i, i+1, \ldots, i+n$, respectively, so after a suitable normalization (or averaging process cf. [4, p. 986]) c_0, c_1,..., c_n are obtained. Henderson has shown that this procedure will always work when D has full rank. A slightly more general (necessary and sufficient) condition for the success of this scheme is given in

THEOREM 6. Let the matrix D of (97) have rank $r+1-\rho$ where $0 \leq \rho \leq r$, and assume that D has the factorization (93) with B as in (99) and with C as in (93) with $c_n \neq 0$. Then there exist $(r+1) \times (r+1)$ matrices P,L,U,S with

P being a permutation matrix, with L being a lower triangular matrix with unit diagonal, with U being an upper triangular matrix with unit diagonal, and with S being a diagonal matrix such that (100) holds if and only if exactly ρ columns of B vanish identically.

Proof: Suppose first that D, C, P, L, U, S are related as in (100). By using (98) in (100) we see that

$$(ULPB - S)\, C = 0$$

and since the last r+1 columns of C are linearly independent this implies that

$$ULPB = S \quad . \tag{102}$$

By hypothesis, the matrices U, L, P are nonsingular so by using (98) and (102) we see that D, B and S must have exactly the same rank $r+1-\rho$ and that B,S have the same null space. In particular, exactly ρ of the diagonal elements s_i of S vanish, the corresponding ρ columns of B must vanish, and (since the rank of B is $r+1-\rho$) no other columns of B can vanish.

Conversely, assume that exactly ρ columns of B vanish (with $r+1-\rho$ being the rank of both D and B). By using elementary row operations we can reduce B to an upper triangular matrix having exactly ρ zeros along the principle diagonal, i.e., we can find P,L such that LPB is an upper

triangular matrix with exactly ρ zero elements along the diagonal. Since each zero column of B is also a zero column of (LP) · B, the ρ columns of LPB which contain the ρ zero pivots must vanish. This being the case, elementary row operations can be used to reduce LPB to a diagonal matrix, i.e., we can find U,S such that (102) holds. Upon multiplying (102) on the right by C and using (98) we then obtain (100). □

NOTE: When D satisfies the conditions of the theorem the matrices LPD and ULPD = SC will be obtained naturally during the two stage elimination process. Indeed, suppose that D can be factored in the form (98) and that $c_0 \neq 0$. (If $c_0 = 0$ we see from (103) and (98) that the first columns of C and D both vanish and that we could replace our problem by one corresponding to a smaller value of r.) Since D has rank $r+1-\rho$, when we use elementary row operations to reduce D to upper echelon form we will end up with exactly ρ zero rows. From (98) we see that each row of D is some linear combination of the rows of C, and since $c_0 \neq 0$ we see that each of the $r+1-\rho$ nonzero rows of our upper echelon matrix must have at least one nonzero element among its first r+1 components. This being the case we can rearrange the ρ zero rows so as to obtain the matrix LPD which has zeros below the principle diagonal and exactly ρ zeros (which lie in the ρ zero rows) along the

principle diagonal. Thus we see that the first stage of the elimination process can be carried out on any matrix D having the factorization (98).

It is the second stage of the elimination process that is dependent upon the auxiliary hypothesis that B has exactly ρ zero columns. Indeed, if the (lower right) r,m - element of LPD is nonzero, then by subtracting suitable multiples of this row from the previous ones we can zero out the upper most r elements of the last column. If this r,m - element is zero, however, we can proceed if and only if the whole last column of LPD is already filled with zeros. Since

$$LPD = (LPB) \cdot C$$

we see that the last column of LPD will vanish if and only if the last column of C is in the null space of LPB and thus in the null space of B, i.e., if and only if the last column of B vanishes. Analogous considerations apply at subsequent stages of the back elimination process.

NOTE. If the matrix D has full rank (as is often the case in practice) then the square matrix B from (98) must also have full rank r+1 so that $\rho=0$ and none of the columns of B vanish. Theorem 6 then guarantees that D has the factorization (100). Moreover, in this case S also has full rank so that U,L,S are uniquely determined by D and the pivoting strategy P.

The following example has been constructed to show that there are cases where the two stage elimination process of Theorem 6 fails even when the roots $z_1,\ldots, z_n$ are uniquely determined by $\mathbf{d}_0, \mathbf{d}_1,\ldots, \mathbf{d}_r$. Indeed, let

$$d_0(z) = 1+2z+z^2+2z^3 = 2(z- 1/2)(z-i)(z+i)$$

$$d_1(z) = \quad z^2 + z^4 = (z-0)(z-0)(z-i)(z+i)$$

$$d_2(z) = 1+2z+2z^2+2z^3+z^4 = (z-1)(z-1)(z-i)(z+i)$$

so that $d_0(z)$, $d_1(z)$, $d_2(z)$ have only the roots $z = \pm i$ in common. In this case

$$B = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 1 & 2 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{bmatrix}$$

and

$$D = BC = \begin{bmatrix} 1 & 2 & 1 & 2 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 2 & 2 & 2 & 1 \end{bmatrix} .$$

The matrices B,D both have rank 2 so that $\rho = 1$, but no column of B vanishes. Upon carrying out the forward elimination process on D we find

$$\begin{bmatrix} 1 & 2 & 1 & 2 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 2 & 2 & 2 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & 1 & 2 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{bmatrix} .$$

The back elimination fails, however, at the point where we try to annihilate the 2 in the second-to-last column of the first row.

8. CONCLUSIONS

Our approach to the problem of transient analysis, i.e., the problem of extracting the system poles $z_1, \ldots, z_n$ from a sequence of samples $\mathbf{y} = (y_0, y_1, y_2, \ldots)$ involves a three stage process. We first assemble a data matrix by applying certain sequence-to-sequence mappings to $\mathbf{y}$ within the conceptual framework of Theorem 3 or Theorem 4. We then compute a null vector for this data matrix by using an eigenvalue analysis or a singular value decomposition (with the former being less costly of computer storage and the latter being somewhat better conditioned.) This null vector then yields a characteristic function having $z_1, \ldots, z_n$ as roots. Our scheme is a conceptually simple one which admits significant new generalizations (such as those of Examples 4 and 7 in Section 6), and it places the existing algorithms within a common mathematical framework.

The sequence-to-sequence mappings which lie at the heart of our analysis can be effected quite simply on a digital computer. We have analyzed the noise suppressing properties of such mappings and identified a fundamental trade off between the size and shape of the region of amplification and the signal-to-noise ratio improvement which can be achieved. The analysis clearly shows why Jain's method is successful in filtering noise from the low order poles when high sampling rates are used (i.e., when the first

few z_ν's are close to 1) and points the way to other preprocessing schemes which will facilitate the computation of the higher order poles. Further noise reduction could be achieved by using an adaptive scheme which first estimates the system poles and then carefully computes them using preprocessing operators which are optimal with respect to the expected pole pattern.

Finally, the problem of estimating the system order n and assessing the accuracy of the computed poles $z_1,\ldots, z_n$ can also be solved by using pole preserving mappings. In the absence of noise the given sequence $\mathbf{y}$ and the auxiliary sequences $\mathbf{y}_0 = \gamma_0(E)\mathbf{y}$, $\mathbf{y}_1 = \gamma_1(E)\mathbf{y},\ldots, \mathbf{y}_r = \gamma_r(E)\mathbf{y}$ will all have exactly the same poles $z_1,\ldots, z_n$ with n unknown. We can perform separate computations of n and the z_ν's using each of the $\mathbf{y}_k$'s, or we can compute approximate null vectors $\mathbf{d}_0, \mathbf{d}_1,\ldots, \mathbf{d}_\nu$ for data matrices constructed using $\mathbf{y}_0, \mathbf{y}_1,\ldots, \mathbf{y}_r$ and then extract n and the z_ν's from the $\mathbf{d}_k$'s by using Henderson's method.

Many of these concepts have been tested by performing the related computations on simple examples, and the results have been most encouraging. A good deal more work will be required, however, in order to perfect these ideas and to incorporate them into efficient production codes.

9. APPENDIX - THE DAMPED VIBRATING STRING

Problems in transient analysis arise when a physical system reverberates in response to some initial excitation. In principle, such a phenomenon can usually be modeled by solving a certain boundary value problem (which characterizes the system) subject to initial conditions which depend upon the form of the initial stimulus. In practice, such models can be analyzed in detail only in extremely simple situations where there is unusual symmetry or low dimensionality. The damped vibrating string provides us with a simple pnysical phenomenon which nicely serves to illustrate the way a problem in transient analysis arises and the inherent difficulties associated with its solution.

Let $u(x,t)$ give the (one dimensional) displacement from the equilibrium position of the string at coordinate x, $0 \leq x \leq L$, at time $t \geq 0$. The motion of the string is governed by the partial differential equation

$$(103)\quad \tau\, u_{xx}(x,t) = \rho\, u_{tt}(x,t) + \varkappa\, u_t(x,t),\quad 0 < x < L,\ t > 0$$

where τ is the tension of the string, ρ is the linear mass density, and $\varkappa$ is the damping coefficient. We shall assume the endpoints of the string are fixed by the boundary conditions

$$(104)\quad u(0,t) = u(L,t) = 0,\quad t \geq 0,$$

and that some external stimulus has subjected the string to the initial displacement and velocity

$$(105)\quad \begin{aligned} u(x,0) &= u_0(x), \quad 0 < x < L \\ u_t(x,0) &= v_0(x), \quad 0 < x < L \end{aligned}$$

at time t=0. Upon separating variables we find that any solution of (103) - (104) has the form

$$(106)\quad u(x,t) = \sum_{m=1}^{\infty} \left\{ A_m e^{(-\alpha+i\omega_m)t} + \bar{A}_m e^{(-\alpha-i\omega_m)t} \right\} \sin(m\pi x/L)$$

where $i^2 = -1$ and where

$$\alpha = \kappa/2\rho$$

$$\omega_m = \left\{ \frac{m^2\pi^2\tau}{L^2\rho} - \frac{\kappa^2}{4\rho^2} \right\}^{1/2}$$

are given in terms of the physical parameters τ, ρ, κ, L which characterize the system. When (106) is subjected to the initial conditions (105) we find that the complex coefficients A_1, A_2, ... are given by the integrals

$$A_m = (\omega_m L)^{-1} \int_{x=0}^{L} \left\{ \omega_m u_0(x) - i[\alpha u_0(x) + v_0(x)] \right\} \sin(m\pi x/L)\,dx, \quad m=1,2,\ldots$$

and thus depend on the choice of $u_0(x)$, $v_0(x)$.

Suppose now that

$$Y(t) = u(x_0,t), \quad t \geq 0$$

is the resulting displacement of the string at some fixed point x_0, $0 < x_0 < L$. Using (106) we see that the transient $Y(t)$ can be written in the form

$$Y(t) = \sum_{m=1}^{\infty} \{[A_m \sin(m\pi x_0/L)]e^{(-\alpha+i\omega_m)t} + [\bar{A}_m \sin(m\pi x_0/L)]e^{(-\alpha-i\omega_m)t}\},$$

or equivalently in the form

$$Y(t) = \sum_{\nu=1}^{\infty} a_\nu e^{s_\nu t}$$

where

$$a_\nu = A_m \sin(m\pi x_0/L) \text{ and } s_\nu = -\alpha+i\omega_m \text{ if } \nu=2m-1,\ m=1,2,\ldots$$

$$a_\nu = \bar{A}_m \sin(m\pi x_0/L) \text{ and } s_\nu = -\alpha-i\omega_m \text{ if } \nu=2m,\ m=1,2,\ldots .$$

The poles s_ν are more or less regularly spaced along the vertical line $\mathrm{Re}\, z = -\alpha$ in the left half plane, cf. Fig. 1, and they depend only on the physical parameters τ, ρ, κ, L of the system. On the other hand, the coefficients a_ν depend on the point of observation x_0 and the initial excitation (i.e., $u_0(x)$ and $v_0(x)$.) The problem of transient analysis is then to recover the s_ν's from certain samples

$$y_k = Y(kT),\ k=0,1,2,\ldots$$

of the signal $Y(t)$.

The intrinsic difficulty of the problem is now apparent. At best we can hope to find the first few of the infinitely

many system poles, s_ν . Moreover, if the initial stimulation of the string fails to excite a given mode, if the point of observation x_0 happens to lie at a node of that mode, or if we unwittingly choose too small a sampling rate, then the mode will be weakly represented (if at all) in the sequence of samples $\mathbf{y} = (y_0, y_1, y_2, \ldots)$ and we will fail to find the corresponding s_ν . Nevertheless, in practice we find that if we use a reasonable sampling rate then it is possible to extract at least the first few s_ν's for "almost" all choices of x_0, $u_0(x)$, and $v_0(x)$. Analogous considerations apply when we use transient analysis to study more complex physical systems which cannot be subjected to such a detailed analysis.

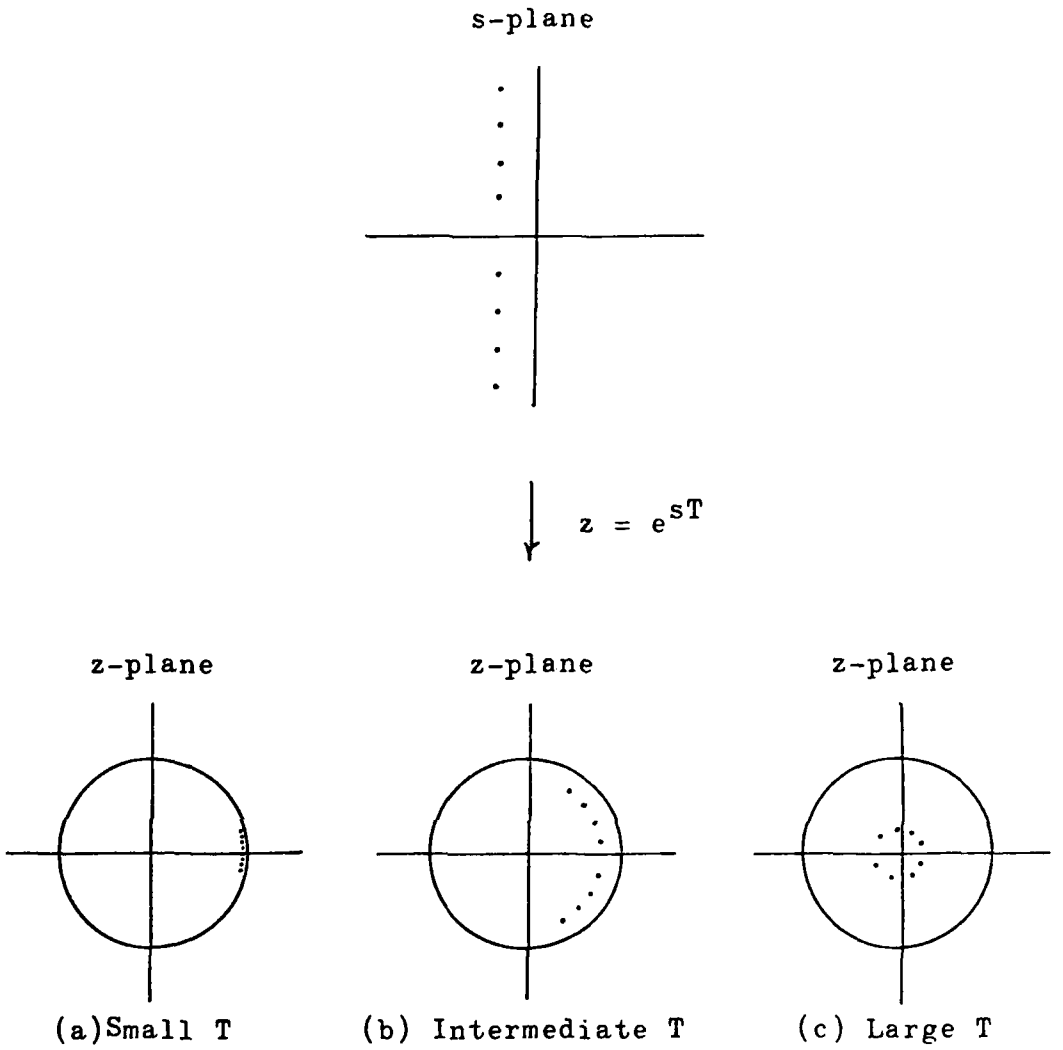


Figure 1. The transformation $z=\exp(sT)$ of (4) maps the s_ν's from the left half s-plane onto the z_ν's within the unit circle of the z-plane, with $T > 0$ being the sampling interval.

(a) When T is too small, the z_ν's are tightly clustered near $z=1$.

(b) When T is correctly chosen, the first few z_ν's are nicely separated.

(c) When T is too large, the z_ν's are burried deep within the unit circle.

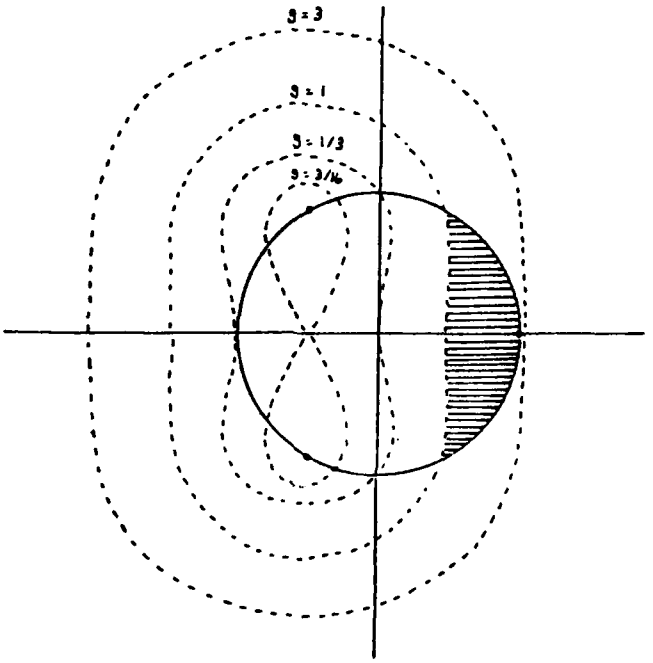


Figure 2. Contour plots of the SNR amplification factor $g(z) = |\ 1+z+z^2\ |^2/3$ for the case where $\gamma(z) = 1+z+z^2$. The region of amplification $\mathcal{A}$ is the set of points z within the unit circle for wnich $g(z) \geq 1$.

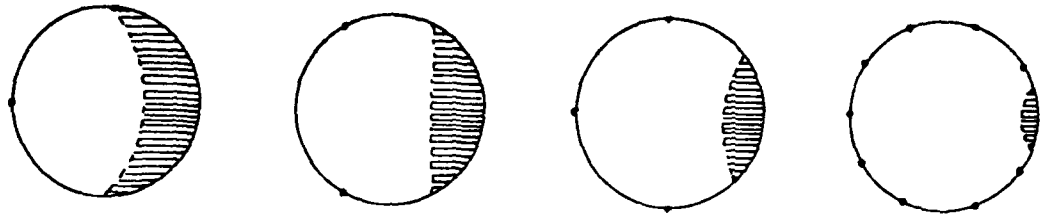


Figure 3. The region of amplification $\mathcal{A}$ for the case where $\gamma(z)=1+z+ \ldots +z^{N-1}$ and N=2,3,4,10. As N increases, the maximum SNR amplification $g(1)=N$ also increases but $\mathcal{A}$ shrinks in size.

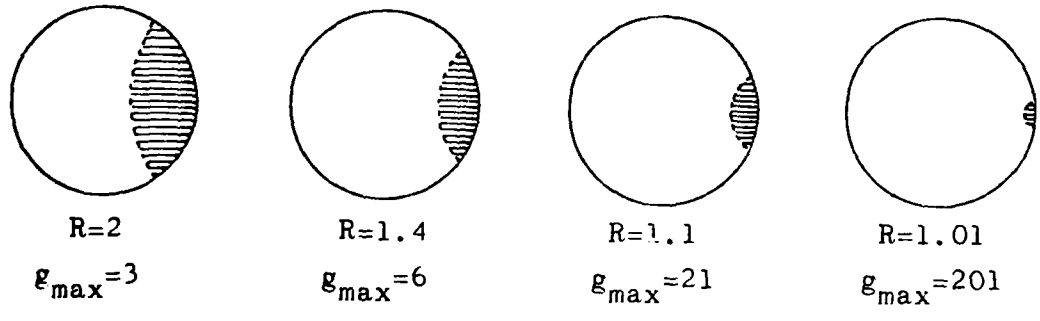


Figure 4. The region of amplification for the case where $\gamma(z) = 1+z/R + z^2/R^2 + \ldots$ and R=2, 1.4, 1.1, 1.01. As $R \downarrow 1$, the maximum SNR amplification $g(1)=(R+1)/(R-1)$ increases but $\mathcal{A}$ shrinks in size.

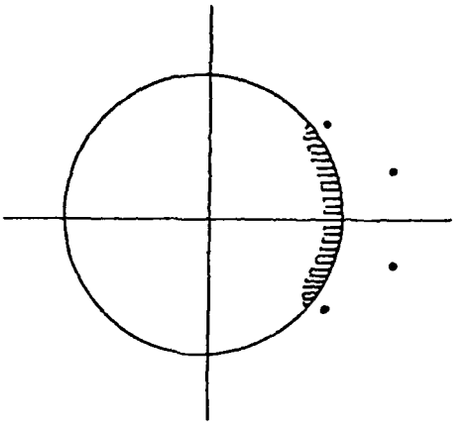


Figure 5. The region of amplification for the case where $\gamma(z)$ is given by (74) with p=2 and with the four poles $(R_i, \pm\theta_i)$, i = 1,2 as shown. In this case α has the approximate form (71) and the corresponding preprocessing scheme is well suited for pole patterns of the form shown in Fig. 1 (b).

REFERENCES

*1. J.R. Auton and M.L. Van Blaricum, Investigation of Procedures for Automatic Resonance Extraction from Noisy Transient Electromagnetic Data, Volumes 1,II,III, Effects Technology, Inc., Santa Barbara, CA., Aug.,1981.

2. J.J. Dongarra, et.al, Linpack User's Guide, SIAM, Philadelphia, PA., 1979.

3. G.E. Forsythe, M.A. Malcolm, and C.B. Moler, Computer Methods for Mathematical Computations, Prentice-Hall, N.Y., 1977.

4. T.L. Henderson, Geometric methods for determining system poles from transient response, IEEE Trans. Acoust., Speech, and Signal Processing, vol. ASSP-29(1981),pp.982-988.

5. V.K. Jain, Filter analysis by use of pencil of functions, IEEE Trans. Circuits Syst. Vol.CAS-21(1974), pp.574-583.

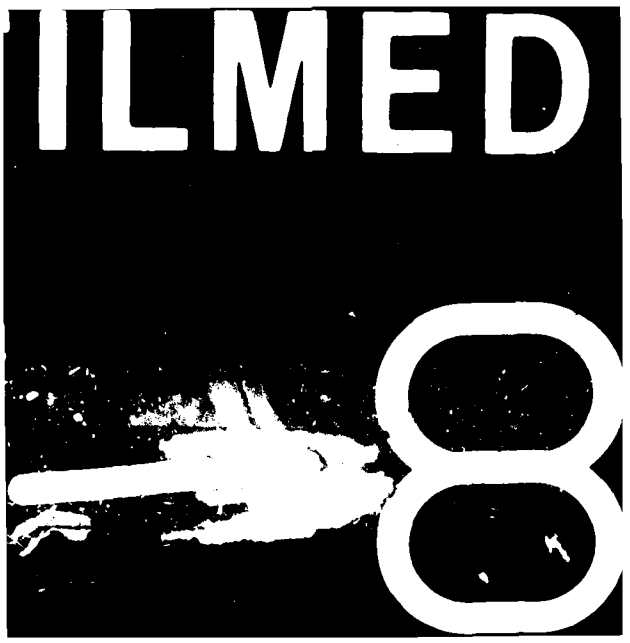
6. T.K. Sarkar, J. Nebat, and V.K. Jain, Suboptimal approximation/identification of transient wave forms from electromagnetic systems by pencil-of-function method, IEEE Trans. Antennas Propagat. Vol.AP-28(1980),pp.928-933.

7. B.T. Smith, et.al., Matrix Eigensystem Routines - EISPACK Guide, Springer-Verlag, N.Y. 1976.

8. G. Strang, Linear Algebra and Its Applications, Second Ed., Academic Press, N.Y., 1980.

9. M.L. Van Blaricum and R. Mittra, Problems and solutions associated with Prony's method for processing transient data, IEEE Trans. Antennas Propagat., Vol.AP-26(1978), pp. 174-182.

* Volume III of this technical report contains an unusually complete annotated bibliography of papers and technical reports dealing with various aspects of the transient analysis problem.


ILMED
8